# Mutational landscape of Hodgkin lymphoma

## Functional consequences and pathogenetic relevance

Fazlyn Reeny Abdul Razak

Cover design: Farah Reeza Abdul Razak
Layout: Fazlyn Reeny Abdul Razak
Printed by: IPSKAMP Drukkers, Enschede, The Netherlands
ISBN (printed): 978-90-367-9284-4
ISBN (digital): 978-90-367-9283-7

**This PhD project was financially supported by:**
Graduate School of Medical Sciences, University Medical Center Groningen
Abel Tasman Talent Program (ATTP)
J.K. de Cock Stichting

**The printing of this thesis was kindly supported by:**
University of Groningen
Graduate School of Medical Sciences, University Medical Center Groningen
Faculty of Medical Sciences, University Medical Center Groningen

university of groningen

# Mutational landscape of Hodgkin lymphoma

Functional consequences and pathogenetic relevance

**PhD thesis**

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. E. Sterken
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Wednesday 16 November 2016 at 09.00 hours

by

**Fazlyn Reeny Binti Abdul Razak**

born on 22 May 1985
in Terengganu, Malaysia

**Supervisor**
Prof. J.H.M. van den Berg

**Co-supervisor**
Dr. A. Diepstra

**Assessment Committee**
Prof. G.A. Huls
Prof. S. de Jong
Prof. P.G. Murray

**Paranymphs**

Miss Rianne Veenstra

Miss Nurul Nadiah Hamidon

I am a cancer cell.

Malicious, relentless, insidious I am, full of destruction, ripe for carnage, yet in my enduring frame I carry the secret of life.

Study me, and you will bring into the light of day precious urns of wisdom long buried in the tomb of ignorance.

Study me, know me, and you will hold the world in fief.

Neglect me, and as surely as the fingers of the dawn grasp first the temples of the East, I will strike you dead.

Fundamentals of Oncology, Henry C. Pitot

# CONTENTS

# Chapter 1

## Introduction and scope of the thesis

**Fazlyn Reeny Abdul Razak**, Arjan Diepstra, Anke van den Berg

**Hodgkin lymphoma**

Lymphoma is a common hematological malignancy in the Western world with about 20 new cases per 100,000 people per year. Although lymphoma classification is dynamic and several revisions have been published, two major categories have always been recognized: Hodgkin lymphoma (HL, ~15% of all cases) and a large group of other lymphomas collectively referred to as non-Hodgkin lymphoma (NHL). In the latest revision of the World Health Organization (WHO) classification of 2016, 61 types of NHL are recognized and diffuse large B cell (DLBCL), follicular lymphoma (FL) and small cell lymphocytic lymphoma are the most common subtypes. Most NHLs are derived from B cells. Figure 1 shows the cell of origin of some of the B cell lymphoma subtypes. Most human B cell NHL and HL derive from germinal center (GC) B cells.

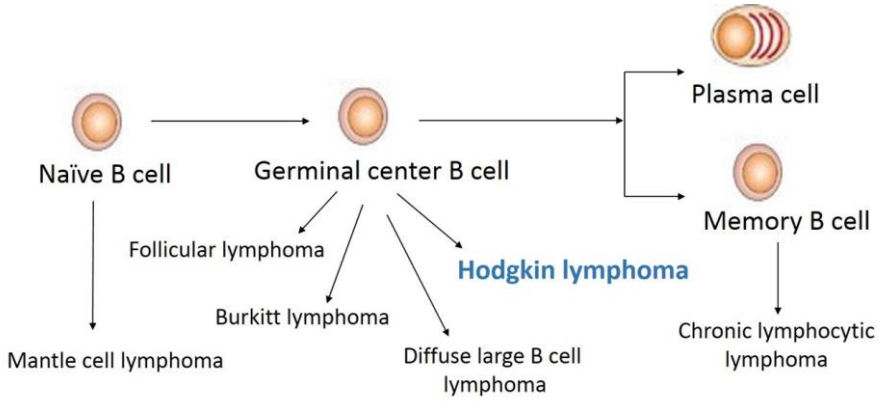



**Figure 1: Cell of origin of different types of B cell lymphoma.** Naïve B cells can differentiate into plasma cell and/or memory B cell in the germinal center upon binding to an antigen. During this process B cells undergo somatic hyper mutations and class switch recombination of the immunoglobulin gene, which presents a high-risk environment for the development of B cell lymphoma. (Adapted from Weigert et al. 2012[1])

HL, formerly known as Hodgkin's disease, was first described by Thomas Hodgkin in 1832.[2] HL is one of the most common lymphomas in the Western world. The incidence varies internationally: 2.3 and 1.9 per 100,000 males and females in more developed regions to 0.8 and 0.5 per 100,000 males and females in less developed regions in 2012.[3] It has a bimodal incidence curve, occurring most frequently in young adulthood (age 15-35) and in individuals over 55 years of age.[4] The WHO classification defines two clinical, pathological

and biological distinct disease entities: classical HL (cHL) accounting for about 95% of all cases and nodular lymphocyte predominant HL (NLPHL).

**Classical Hodgkin lymphoma**

Classical HL is divided into nodular sclerosis (NS), mixed cellularity (MC), lymphocyte depleted (LD) and lymphocyte rich (LR) subtypes based on the morphology of the tumor cells and the composition of the background.[5] NS is the most common histological subtype, followed by MC, comprising approximately 75% and 15% of cHL cases respectively. The other cHL subtypes are rare.[6] cHL is characterized by the presence of large mono- or multinucleated Hodgkin cells and binucleated Reed-Sternberg (HRS) cells which were first described by Dorothy Reed and Carl Sternberg around 1900. The HRS cells usually account for only about 1% of the total cell number in the tumor mass. The majority of the cells present in the involved lymph nodes are non-neoplastic B and T cells, macrophages, eosinophils, neutrophils, plasma cells and fibroblasts.[7,8] Although HRS cells are derived from GC B cells, they have lost their B cell phenotype almost completely. Table 1 shows the genetic and phenotypic characteristics of HRS cells in relation to their cellular origin.

HRS cells lack expression of the B cell receptor (BCR) in most cases and rely on activation of multiple pathways to escape from apoptosis that is normally induced in GC B cells that lack a functional BCR. Two crucial signaling pathways are activated in HRS cells, i.e. the NF-κB and the JAK/STAT pathways. Activation of both the canonical and alternative NF-κB pathway is induced by CD30, CD40-CD40L and Epstein Barr virus (EBV)-derived latent membrane protein 1 (LMP1). The JAK/STAT pathway involves a complex network of cytokine signaling events leading to the activation and nuclear translocation of STAT transcription factors.[10]

**Table 1**: Genetic and Phenotype characteristics of HRS cells

| Putative cell of origin | Defective, pre-apoptotic GCB cell |
|---|---|
| Phenotype | |
| CD30 expression | Yes |
| CD15 expression | Yes (~70%) |
| B cell receptor expression | No |
| Loss of most B cell markers | Yes |
| Expression of GCB cell markers (BCL6, AID) | Rarely |
| Expression of markers for non-B cells | |
| CD3 and granzyme B | Infrequently |
| TARC and CD15 | Often |
| Expression of plasma markers (Mum1, CD38) | Often |
| Expression of B-lineage commitment factor Pax-5 | Yes (low) |
| EBV positivity | Yes (~40%) |
| Signaling pathways | |
| NFKB activation | Yes |
| JAK/STAT activation | Yes |
| Genetic lesions | |
| Aberrant expression of multiple RTKs | Yes (~60-100%) |
| NFKBIA mutations | Yes (~10-20%) |
| NFKBIE mutations | Yes (~10%) |
| TNFAIP3 mutations | Yes (~40%) |
| REL gains/amplification | Yes (~50%) |
| MAP3K14 (NIK) gains/amplifications | Yes (~25%) |
| JAK2, PD-L1, PD-L2, JMJD2C gains/amplification | Yes (~30%) |
| SOCS1 mutations | Yes (~40%) |
| MHC2TA translocation | Yes (15%) |

(Adapted from Rosenwald & Kuppers 2015[9])

EBV has been detected in HRS cells in a proportion of the cHL patients.[11] In the western world, up to 40% of cHL cases are associated with EBV and the frequency is highest in the MC subtype. EBV displays a latency type II infection pattern limiting expression to LMP1, LMP2A, EBV nuclear antigen 1 (EBNA1) and two small RNAs (EBER1 and EBER2).[12] In EBV+ cases, the viral genome is clonally present in all HRS cells. EBV plays a central role in cHL pathogenesis by rescuing BCR deficient, pre-apoptotic GC B cells from apoptosis, thus rendering independence of survival signals normally supplied by the BCR.[13] LMP1 mimics activated CD40 and contribute to the activation of

NF-κB. LMP2A mimics the BCR and provides anti-apoptotic signals to the infected B cells.[14] Tumor cell EBV positivity is associated with treatment failure in older adult cHL patients.[15]

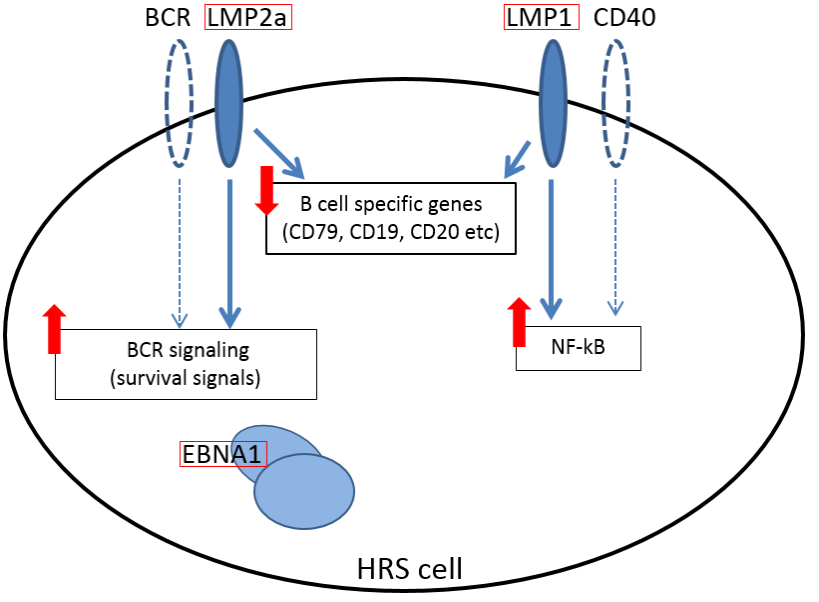



**Figure 2: The role of EBV latency type II proteins in malignant transformation and survival of HRS cells.** LMP1 mimics activated CD40 and contribute to the activation of NF-κB. LMP2A mimics the BCR and provides anti-apoptotic signals to the infected B cells. (Adapted from Massini et al. 2009[16])

HLA class I expression is observed in HRS cells of 70% of the EBV+ cHL and in 14% of the EBV- cHL cases.[17] Studies from our group[18] and Hjalgrim et al.[19] showed that increased risk of EBV+ cHL was associated with HLA-A*01 and decreased risk with HLA-A*02. The effects of HLA-A*01 and HLA-A*02 were independent of each other but dependent on allele copy number, such that HLA-A*01 homozygotes had an almost 10-fold greater odds of EBV-associated HL than HLA-A*02 homozygotes. It is known that HLA-A*01 alleles are less efficient in presenting EBV latency type II derived antigenic peptides, whereas HLA-A*02 can efficiently present these peptides and induce effective cytotoxic T cell responses.[20,21] Expression of HLA class II is observed in about 50% of cHL cases, irrespective of EBV status. Down regulation of HLA class II by HRS cells is an adverse prognostic factor.[22] Genome-wide association study (GWAS) in cHL patients showed the most significant association with

rs6903608. This single nucleotide polymorphism was located within the HLA class II region.[23] Other susceptibility loci, i.e. at 2p16.1 (REL), 8q24.21 (PVT1), 10p14 (GATA3) and 19p13.3 (TCF3) were also identified by GWAS.[23-25]

**Common genetic aberrations in Hodgkin lymphoma**

***Copy number variations***

Genomic instability is a characteristic feature of HRS cells.[26] HRS cells frequently harbor recurrent numerical and structural aberrations as detected by classical cytogenetic and fluorescence in situ hybridization analysis.[27] In contrast to NHL that often harbor subtype specific chromosomal translocations, recurrent chromosomal translocations in cHL are less common.[28] However, results from molecular genetic studies using comparative genomic hybridization and allelotyping do indicate typical copy number variations.[26] Chromosomal gains most frequently involve chromosome arms 2p, 9p, 16p, and 17q and losses involve 13q, 6q, and 11q.[29-31] Steidl et al.[32] showed that gains of chromosome 16p are associated with treatment failure. The *REL* locus (2p15-16) shows genomic gains and amplifications in nearly half of cHL cases and these alterations correlate with REL protein levels.[29,33] The *MAP3K14* (also called NIK) gene at 17q21.31, encoding a main factor of the alternative NF-κB pathway, shows copy number gains in 31% of the cHL cases.[32,34] The chromosomal region 9p24, which includes the *JAK2* gene locus, shows genomic gains in 30% of the cHL.[35,36] Comparative genomic hybridization (CGH) of microdissected HRS cells from 10 cHL cases showed gains of *STAT6* (12q13), *NOTCH1* (9q34) and *JUNB* (19p13).[31] In addition to these common gains and losses of specific chromosomal regions, there is one recurrent chromosomal translocation, involving the MHC class II transactivator (CIITA) gene locus that has been observed in 15% of the cHL cases.[28] ABCC1, the multidrug resistance gene also shows genomic gains and overexpression in a cHL cell line, and the overexpression contributed to the drug resistance phenotype.[32] Salipante et al.[37] observed copy number loss of the candidate tumor suppressor TNFRSF14 locus in 8 out of 19 cases using low-coverage whole genome sequencing of 500 flow sorted HRS cells per case. PDL1 copy number gain was reported to correlate with increased PDL1 expression in HL. Ansell et al.[38] showed that concurrent gain of the PDL1 and PDL2 loci, increased expression of both PD-1 ligands, and evidence of active JAK-STAT signaling.

***Gene mutations***

As constitutive activation of the NF-κB and JAK/STAT pathways have been linked to cHL pathogenesis, most of the targeted studies focused on the genes involved in these pathways. Initially studies were performed in the available HL cell lines. In case gene mutations were found in the cell lines, microdissected primary HRS cells were analyzed for confirmation.

NF-κB consists of a family of transcription factors that play critical roles in inflammation, immunity, cell proliferation, differentiation, and survival.[39] The NF-κB transcription factor family consists of five members, RelA (p65), RelB, c-Rel, NF-κB1 (p50/p105), and NF-κB2 (p52/p100), which can act as homo- and/or heterodimers.[40] Multiple somatic mutations were found in members of this pathway (Figure 2). In addition, somatic mutations were observed in two inhibitors of the NF-κB signaling pathway, *NFKB1A* and *NFKB1E*, in ~20% of cHL cases. These two factors bind to NF-κB in the cytoplasm and prevent their nuclear translocation.[41] Other genetic abnormalities affecting the NF-κB pathway include mutations in the TNFα-induced protein 3 (*TNFAIP3*) gene, which encodes A20, a negative regulator of NF-κB. These loss-of-function mutations are found in 40% of EBV-negative HL and provide HRS cells with an alternative mechanism to escape from apoptosis.[41] In rare cases, mutations of *CYLD* or *TRAF3*, which can positively regulate NF-κB activity, are detected in HRS cells. [34,42]

The JAK/STAT pathway consists of a complex network of cytokine signaling transduction molecules, which lead to the activation and nuclear translocation of STAT transcription factors.[43] A variety of cytokines and growth factors can bind to their corresponding receptors on the cell surface of HRS cells and activate JAK. Activated JAK induces phosphorylation of the STAT proteins on specific tyrosine residues. Previous studies showed that STAT3, STAT5a, STAT5b and STAT6 are highly active in HRS cells.[44-46] SOCS1, encoding the main inhibitor of JAK/STAT signaling, is inactivated by mutations in approximately 40% of the cHL cases.[47] PTPN1, another negative regulator of the JAK/STAT signaling pathway is mutated in 20% of the HL cases.[48]

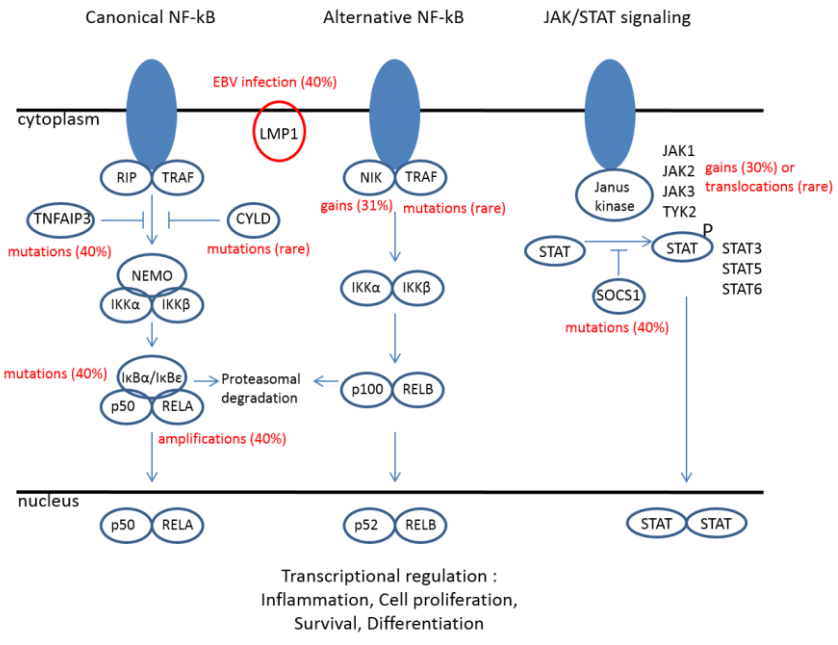



**Figure 3: Common aberrations in NF-κB and JAK/STAT signaling pathway genes observed in HRS cells.** NF-κB is activated through two signaling cascades: the canonical and alternative pathways. Canonical NF-κB pathway of NF-κB activation relies on inducible degradation of IκBs, particularly IκBα, leading to nuclear translocation of various NF-κB complexes, predominantly the p50/RelA dimer. As for alternative NF-κB pathway, it activates the RelB/p52 NF-κB complex using a mechanism that relies on the inducible processing of p100 instead of degradation of IκBα. HRS cells have constitutive activity of both NF-κB pathways. The JAK/STAT pathway is the main signaling pathway for cytokines. The frequency of genetic lesions and viral infections affecting NF-κB or STAT activity in HRS cells is indicated as percentages. (Adapted from Steidl et al.[48] & Kuppers et al.[10])

### Next generation sequencing approaches applied to purified HRS cells

Most genomic studies of cHL are based on cHL cell lines and nested PCR approaches in HRS cells and were focused on genes involved in NF-kB and JAK/STAT pathways. A disadvantage of this method is that it does not allow identification of new targets. Next generation sequencing (NGS) technologies have provided numerous opportunities to comprehensively screen the entire genome for aberrations in cancer cell.[50,51] Strategies for investigating the cancer genome include sequencing the whole genome (WGS) or limiting the sequencing to exome (WES), transcriptome (RNA-seq) or targeted sequencing of specific exons. Previous genomic studies of cHL have largely been confined

to cell lines due to the difficulty of isolating sparse HRS cells from the reactive background cells in sufficient numbers. As a consequence, there are only a limited number of published studies that use NGS technology in HL.

Steidl et al.[28] applied RNA-seq to study 2 HL cell lines as a discovery platform. They found a gene fusion involving the transcription factor CIITA located on chromosome 16p and PD1-ligand locus on chromosome 9p in one cell line. Analysis of additional samples indicated CIITA translocations in 15% of cHL cases, with promiscuous translocation partners. The same group[52] also performed WGS in L428 and RNA-seq in L1236 and DEV. By using genomic breakpoint and fusion transcript predicting algorithms, they identified another 2 novel fusion transcripts within 9p24.1 locus, KIAA1432-CLDN14 (L428) and PDCD1LG2-IGHV7-81 (L1236).

Two different approaches to purify HRS cells have been applied so far. One approach is to microdissect HRS cells from fresh-frozen tissue biopsies by laser dissection microscopy (LDM). Hartmann et al.[31] used DNA from approximately 100,000 microdissected HRS cells from 12 primary cHL cases with at least 70% HRS cells in the microdissected areas for aCGH. They showed that genomic regions with recurrent copy number aberrations in HRS cells include genes constitutively expressed in cHL. Steidl et al.[32] reported copy number alterations in more than 20% of cases including gains of 2p, 9p, 16p, 17q, 19q, 20q, and losses of 6q, 11q, and 13q by genome wide copy number analysis from pools of 500 to 1,000 HRS cells per case with whole genome amplification (WGA). Using RNA isolated from approximately 1,000 HRS cells per case, Steidl et al.[53] found significant overlap of the gene expression profiles of primary HRS cells and HL cell lines. Tiacci et al.[54] showed similar overall levels of hallmark cHL gene signatures in pooled of 1,000 to 2,000 microdissected HRS cells per cases from frozen biopsies and cHL cell lines by microarray analysis. Karube et al.[55] reported an expression profile of 140 genes of chemokines, cytokines and their receptors in microdissected HRS cells from 14 HL tissue samples. Another way to purify HRS cells is by flow cytometry of cHL cell suspensions using a mixture of antibodies (CD15, CD86, CD45, CD40, and CD30).[56,57] In a recent study, Reichel et al.[58] isolated 1,000 to 100,000 HRS cells using an antibody cocktail consisting of CD64, CD30, CD5, CD40, CD20, CD15, CD45 and CD95 for whole exome sequencing and independently validated a selection of variants by RNA-seq. In this study they reported that B2M was the most frequently mutated gene in flow purified primary HRS cells.

**Scope of the thesis**

The aim of this thesis was to identify additional mutations in HL cell lines by a comprehensive sequencing approach and to substantiate the functional importance of some of these mutations in HL. In **chapter 2,** we performed WES on 7 HL cell lines to provide a comprehensive overview of the mutational landscape in HL. In **chapter 3**, we focused on mutations in CD58 and studied its expression in HL cell lines and primary HL tissue samples. In **chapter 4**, we studied the possible involvement of CSF2RB in the pathogenesis of HL. In **chapter 5**, we determined the role of wild type and mutant *MYB* in HL cell lines by determining effects on tumor cell growth and by defining the set of MYB regulated genes. Finally, in **chapter 6** we summarize and discuss the findings of this thesis and place our findings in perspective for future research.

## REFERENCES


1. Weigert O, Weinstock DM. The evolving contribution of hematopoietic progenitor cells to lymphomagenesis. Blood 2012; 120(13): 2553-61.
2. Hodgkin. On some Morbid Appearances of the Absorbent Glands and Spleen. Med Chir Trans 1832; 17: 68-114.
3. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013; 49(6): 1374-403.
4. Mehta V, Sarda A, Balachandran C, Rao R, Monga P. Late onset nodular prurigo - the sole and initial manifestation of occult Hodgkin's disease. Indian J Dermatol 2009; 54(2): 192-3.
5. Eberle FC, Mani H, Jaffe ES. Histopathology of Hodgkin's lymphoma. Cancer J 2009; 15(2): 129-37.
6. Swerdlow SH, Jaffe ES, International Agency for Research on Cancer, World Health Organization. WHO classification of tumours of haematopoietic and lymphoid tissues, International Agency for Research on Cancer: Lyon, 2008.
7. Jaffe ES, World Health Organization. International Agency for Research on Cancer. Pathology and genetics of tumours of haematopoietic and lymphoid tissues, IARC Press;: Lyon, 2001.
8. Pileri SA, Ascani S, Leoncini L, Sabattini E, Zinzani PL, Piccaluga PP et al. Hodgkin's lymphoma: the pathologist's viewpoint. J Clin Pathol 2002; 55(3): 162-76.
9. Küpper ARR. Pathology and molecular pathology of Hodgkin lymphoma. In: Hodgkin lymphoma. Springer International Publishing: Berlin Heidelberg, 2015, pp 45-78.
10. Kuppers R, Engert A, Hansmann ML. Hodgkin lymphoma. J Clin Invest 2012; 122(10): 3439-47.
11. Gulley ML, Eagan PA, Quintanilla-Martinez L, Picado AL, Smir BN, Childs C et al. Epstein-Barr virus DNA is abundant and monoclonal in the Reed-Sternberg cells of Hodgkin's disease: association with mixed cellularity subtype and Hispanic American ethnicity. Blood 1994; 83(6): 1595-602.
12. Garcia-Cosio M, Santon A, Martin P, Reguero ME, Cristobal E, Bellas C. Analysis of Epstein-Barr virus strains and variants in classical Hodgkin's lymphoma by laser microdissection. Histol Histopathol 2008; 23(2): 209-17.
13. Bechtel D, Kurth J, Unkel C, Kuppers R. Transformation of BCR-deficient germinal-center B cells by EBV supports a major role of the virus in the pathogenesis of Hodgkin and posttransplantation lymphomas. Blood 2005; 106(13): 4345-50.
14. Ersing I, Bernhardt K, Gewurz BE. NF-kappaB and IRF7 pathway activation by Epstein-Barr virus Latent Membrane Protein 1. Viruses 2013; 5(6): 1587-606.
15. Diepstra A, van Imhoff GW, Schaapveld M, Karim-Kos H, van den Berg A, Vellenga E et al. Latent Epstein-Barr virus infection of tumor cells in classical Hodgkin's lymphoma predicts adverse outcome in older adult patients. J Clin Oncol 2009; 27(23): 3815-21.
16. Massini G, Siemer D, Hohaus S. EBV in Hodgkin Lymphoma. Mediterr J Hematol Infect Dis 2009; 1(2): e2009013.
17. Niens M, van den Berg A, Diepstra A, Nolte IM, van der Steege G, Gallagher A et al. The human leukocyte antigen class I region is associated with EBV-positive Hodgkin's lymphoma: HLA-A and HLA complex group 9 are putative candidate genes. Cancer Epidemiol Biomarkers Prev 2006; 15(11): 2280-4.
18. Niens M, Jarrett RF, Hepkema B, Nolte IM, Diepstra A, Platteel M et al. HLA-A*02 is associated with a reduced risk and HLA-A*01 with an increased risk of developing EBV+ Hodgkin lymphoma. Blood 2007; 110(9): 3310-5.

**19.** Hjalgrim H, Rostgaard K, Johnson PC, Lake A, Shield L, Little AM et al. HLA-A alleles and infectious mononucleosis suggest a critical role for cytotoxic T-cell response in EBV-related Hodgkin lymphoma. Proc Natl Acad Sci U S A 2010; 107(14): 6400-5.

**20.** Misko IS, Schmidt C, Honeyman M, Soszynski TD, Sculley TB, Burrows SR et al. Failure of Epstein-Barr virus-specific cytotoxic T lymphocytes to lyse B cells transformed with the B95-8 strain is mapped to an epitope that associates with the HLA-B8 antigen. Clin Exp Immunol 1992; 87(1): 65-70.

**21.** Khanna R, Burrows SR, Kurilla MG, Jacob CA, Misko IS, Sculley TB et al. Localization of Epstein-Barr virus cytotoxic T cell epitopes using recombinant vaccinia: implications for vaccine development. J Exp Med 1992; 176(1): 169-76.

**22.** Diepstra A, van Imhoff GW, Karim-Kos HE, van den Berg A, te Meerman GJ, Niens M et al. HLA class II expression by Hodgkin Reed-Sternberg cells is an independent prognostic factor in classical Hodgkin's lymphoma. J Clin Oncol 2007; 25(21): 3101-8.

**23.** Enciso-Mora V, Broderick P, Ma Y, Jarrett RF, Hjalgrim H, Hemminki K et al. A genome-wide association study of Hodgkin's lymphoma identifies new susceptibility loci at 2p16.1 (REL), 8q24.21 and 10p14 (GATA3). Nat Genet 2010; 42(12): 1126-30.

**24.** Urayama KY, Jarrett RF, Hjalgrim H, Diepstra A, Kamatani Y, Chabrier A et al. Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus status-defined subgroups. J Natl Cancer Inst 2012; 104(3): 240-53.

**25.** Cozen W, Timofeeva MN, Li D, Diepstra A, Hazelett D, Delahaye-Sourdeix M et al. A meta-analysis of Hodgkin lymphoma reveals 19p13.3 TCF3 as a novel susceptibility locus. Nat Commun 2014; 5: 3856.

**26.** Re D, Zander T, Diehl V, Wolf J. Genetic instability in Hodgkin's lymphoma. Ann Oncol 2002; 13 Suppl 1: 19-22.

**27.** Yurchenko M, Sidorenko SP. Hodgkin's lymphoma: the role of cell surface receptors in regulation of tumor cell fate. Exp Oncol 2010; 32(4): 214-23.

**28.** Steidl C, Shah SP, Woolcock BW, Rui L, Kawahara M, Farinha P et al. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. Nature 2011; 471(7338): 377-81.

**29.** Joos S, Menz CK, Wrobel G, Siebert R, Gesk S, Ohl S et al. Classical Hodgkin lymphoma is characterized by recurrent copy number gains of the short arm of chromosome 2. Blood 2002; 99(4): 1381-7.

**30.** Chui DT, Hammond D, Baird M, Shield L, Jackson R, Jarrett RF. Classical Hodgkin lymphoma is associated with frequent gains of 17q. Genes Chromosomes Cancer 2003; 38(2): 126-36.

**31.** Hartmann S, Martin-Subero JI, Gesk S, Husken J, Giefing M, Nagel I et al. Detection of genomic imbalances in microdissected Hodgkin and Reed-Sternberg cells of classical Hodgkin's lymphoma by array-based comparative genomic hybridization. Haematologica 2008; 93(9): 1318-26.

**32.** Steidl C, Telenius A, Shah SP, Farinha P, Barclay L, Boyle M et al. Genome-wide copy number analysis of Hodgkin Reed-Sternberg cells identifies recurrent imbalances with correlations to treatment outcome. Blood 2010; 116(3): 418-27.

**33.** Martin-Subero JI, Gesk S, Harder L, Sonoki T, Tucker PW, Schlegelberger B et al. Recurrent involvement of the REL and BCL11A loci in classical Hodgkin lymphoma. Blood 2002; 99(4): 1474-7.

**34.** Otto C, Giefing M, Massow A, Vater I, Gesk S, Schlesner M et al. Genetic lesions of the TRAF3 and MAP3K14 genes in classical Hodgkin lymphoma. Br J Haematol 2012; 157(6): 702-8.

**35.** Joos S, Granzow M, Holtgreve-Grez H, Siebert R, Harder L, Martin-Subero JI et al. Hodgkin's lymphoma cell lines are characterized by frequent aberrations on chromosomes 2p and 9p including REL and JAK2. Int J Cancer 2003; 103(4): 489-95.

**36.** Joos S, Kupper M, Ohl S, von Bonin F, Mechtersheimer G, Bentz M et al. Genomic imbalances including amplification of the tyrosine kinase gene JAK2 in CD30+ Hodgkin cells. Cancer Res 2000; 60(3): 549-52.

**37.** Salipante SJ, Adey A, Thomas A, Lee C, Liu YJ, Kumar A et al. Recurrent somatic loss of TNFRSF14 in classical Hodgkin lymphoma. Genes Chromosomes Cancer 2016; 55(3): 278-87.

**38.** Ansell SM, Lesokhin AM, Borrello I, Halwani A, Scott EC, Gutierrez M et al. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. N Engl J Med 2015; 372(4): 311-9.

**39.** Oeckinghaus A, Ghosh S. The NF-kappaB family of transcription factors and its regulation. Cold Spring Harb Perspect Biol 2009; 1(4): a000034.

**40.** Lernbecher T, Muller U, Wirth T. Distinct NF-kappa B/Rel transcription factors are responsible for tissue-specific and inducible gene activation. Nature 1993; 365(6448): 767-70.

**41.** Schmitz R, Stanelle J, Hansmann ML, Kuppers R. Pathogenesis of classical and lymphocyte-predominant Hodgkin lymphoma. Annu Rev Pathol 2009; 4: 151-74.

**42.** Schmidt A, Schmitz R, Giefing M, Martin-Subero JI, Gesk S, Vater I et al. Rare occurrence of biallelic CYLD gene mutations in classical Hodgkin lymphoma. Genes Chromosomes Cancer 2010; 49(9): 803-9.

**43.** Kuppers R. Molecular biology of Hodgkin lymphoma. Hematology Am Soc Hematol Educ Program 2009: 491-6.

**44.** Scheeren FA, Diehl SA, Smit LA, Beaumont T, Naspetti M, Bende RJ et al. IL-21 is expressed in Hodgkin lymphoma and activates STAT5: evidence that activated STAT5 is required for Hodgkin lymphomagenesis. Blood 2008; 111(9): 4706-15.

**45.** Baus D, Pfitzner E. Specific function of STAT3, SOCS1, and SOCS3 in the regulation of proliferation and survival of classical Hodgkin lymphoma cells. Int J Cancer 2006; 118(6): 1404-13.

**46.** Skinnider BF, Mak TW. The role of cytokines in classical Hodgkin lymphoma. Blood 2002; 99(12): 4283-97.

**47.** Weniger MA, Melzner I, Menz CK, Wegener S, Bucur AJ, Dorsch K et al. Mutations of the tumor suppressor gene SOCS-1 in classical Hodgkin lymphoma are frequent and associated with nuclear phospho-STAT5 accumulation. Oncogene 2006; 25(18): 2679-84.

**48.** Gunawardana J, Chan FC, Telenius A, Woolcock B, Kridel R, Tan KL et al. Recurrent somatic mutations of PTPN1 in primary mediastinal B cell lymphoma and Hodgkin lymphoma. Nat Genet 2014; 46(4): 329-35.

**49.** Steidl C, Connors JM, Gascoyne RD. Molecular pathogenesis of Hodgkin's lymphoma: increasing evidence of the importance of the microenvironment. J Clin Oncol 2011; 29(14): 1812-26.

**50.** Daniels M, Goh F, Wright CM, Sriram KB, Relan V, Clarke BE et al. Whole genome sequencing for lung cancer. J Thorac Dis; 4(2): 155-63.

**51.** Liang WS, Craig DW, Carpten J, Borad MJ, Demeure MJ, Weiss GJ et al. Genome-wide characterization of pancreatic adenocarcinoma patients using next generation sequencing. PLoS One; 7(10): e43192.

**52.** Twa DD, Chan FC, Ben-Neriah S, Woolcock BW, Mottok A, Tan KL et al. Genomic rearrangements involving programmed death ligands are recurrent in primary mediastinal large B-cell lymphoma. Blood 2014; 123(13): 2062-5.

**53.** Steidl C, Diepstra A, Lee T, Chan FC, Farinha P, Tan K et al. Gene expression profiling of microdissected Hodgkin Reed-Sternberg cells correlates with treatment outcome in classical Hodgkin lymphoma. Blood 2012; 120(17): 3530-40.

**54.** Tiacci E, Doring C, Brune V, van Noesel CJ, Klapper W, Mechtersheimer G et al. Analyzing primary Hodgkin and Reed-Sternberg cells to capture the molecular and cellular pathogenesis of classical Hodgkin lymphoma. Blood 2012; 120(23): 4609-20.

**55.** Karube K, Ohshima K, Suzumiya J, Kawano R, Kikuchi M, Harada M. Gene expression profile of cytokines and chemokines in microdissected primary Hodgkin and Reed-Sternberg (HRS) cells: high expression of interleukin-11 receptor alpha. Ann Oncol 2006; 17(1): 110-6.

**56.** Fromm JR, Kussick SJ, Wood BL. Identification and purification of classical Hodgkin cells from lymph nodes by flow cytometry and flow cytometric cell sorting. Am J Clin Pathol 2006; 126(5): 764-80.

**57.** Salipante SJ, Mealiffe ME, Wechsler J, Krem MM, Liu Y, Namkoong S et al. Mutations in a gene encoding a midbody kelch protein in familial and sporadic classical Hodgkin lymphoma lead to binucleated cells. Proc Natl Acad Sci U S A 2009; 106(35): 14920-5.

**58.** Reichel J, Chadburn A, Rubinstein PG, Giulino-Roth L, Tam W, Liu Y et al. Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. Blood 2015; 125(7): 1061-72.

# Chapter 2

## The mutational landscape of Hodgkin lymphoma cell lines determined by whole exome sequencing

Yuxuan Liu, **Fazlyn Reeny Abdul Razak**, Martijn Terpstra, Fong Chun Chan, Ali Saber, Marcel Nijland, Gustaaf van Imhoff, Lydia Visser, Randy Gascoyne, Christian Steidl, Joost Kluiver, Arjan Diepstra, Klass Kok, Anke van den Berg

Leukemia July 2014 (Doi: 10.1038/leu.2014.201)

**ABSTRACT**

Hodgkin lymphoma (HL) is characterized by constitutive activation of several signaling pathways and transcription factors, which is partly caused by gene mutations. To generate an overview of the mutational landscape in HL we performed whole exome-sequence (WES) analysis of 7 HL cell lines. Overall, we identified 463 genes mutated in 2 or more HL cell lines and 373 genes mutated in 2 or more classical HL cell lines. Based on SNPEFF_IMPACT, PolyPhen2 and SIFT analyses we showed that approximately half of the mutations have a putative damaging effect. As compared to Broad Institute data, an overall consistency ranging from 72.1% to 98.5% was observed. With the exception of the second deletion in *SOCS1* reported in L1236, all mutations were confirmed. We identified mutations in HLA associated genes, *B2M, HLA-A, HLA-DRB1* and in *CIITA*. Consistent with the *B2M* mutations affecting the ATG start codon, we observed no or very low membrane B2M and HLA class I expression in L428 and DEV by flow cytometry. *B2M* mRNA levels were reduced in both cell lines as compared to L1236, whereas HLA-A, HLA-B and HLA-C levels were in the same range. Our WES data revealed *CIITA* mutations in 2 of the 7 HL cell lines albeit with low read counts, that were confirmed by Sanger sequencing and also Broad data set. No significant enrichment of mutated genes in genomic regions with copy number gain or loss was observed. Combining mutation status with mRNA expression levels revealed a differential expression pattern in cHL as compared to germinal center B cells for 44 of the 373 genes mutated in cHL. In conclusion, we observed a high number of consistently mutated genes, with part of them mapping to regions with copy number gain or loss and part of them showing deregulated expression.

## INTRODUCTION

Hodgkin lymphoma (HL) is characterized by a minority of B-cell derived tumor cells, named Hodgkin Reed-Sternberg (HRS) cells in classical (c)HL and lymphocyte predominant (LP) cells in nodular lymphocyte predominant (NLP) HL. HRS cells lack expression of the B-cell receptor in most cases and rely on activation of multiple pathways to escape from apoptosis in the germinal center reaction.[1] Constitutive activation of the NF-κB pathway is achieved by activation of various signaling pathways, e.g. CD30, CD40-CD40L and Epstein Barr virus (EBV)-derived latent membrane protein 1 (LMP1). In addition, NF-κB activation is caused by diverse aberrations in multiple genes involved in the NF-κB pathway, such as amplifications of *REL,* gains of *MAP3K14*, mutations of *TNFAIP3*, *NFKBIA* and *NFKBIE*. A second pathway frequently altered in HRS cells is the JAK/STAT signaling pathway. Activation is caused by gains of *JAK2* and inactivating mutations of *SOCS1*. In recent years, high-throughput sequencing technology has provided novel opportunities for the comprehensive identification of genetic aberrations involved in various types of cancer. To increase our understanding of the pathogenesis of HL, we performed whole exome sequencing (WES) in seven cell lines to identify commonly mutated genes in HL.

## MATERIALS AND METHODS

### *Cell lines and germinal center B cells*

The cHL cell lines L428 (nodular sclerosis), L1236 (mixed cellularity), KMH2 (mixed cellularity), L591 (nodular sclerosis, EBV+) and L540 (nodular sclerosis, T cell derived) were cultured in RPMI 1640 medium (Lonza Walkersville, Walkersville, MD) supplemented with 5% (L428), 20% (DEV) and 10% (other cell lines) fetal calf serum, 100U/ml penicillin/streptomycin and ultraglutamine (Lonza Walkersville) in a 5% CO2 atmosphere at 37°C. SUPHD1 (lymphocyte depleted subtype) was cultured in McCoy 5A medium supplemented with 10% fetal calf serum. Germinal center (GC) B cells were purified from human tonsils based on expression of CD19+IgD-CD38+.

### *Exome sequencing*

Whole exome sequencing (WES) was carried out using standardized protocols of the UMCG genome facility on seven HL cell lines. Briefly, 3µg genomic DNA was randomly fragmented by ultrasound Nebulisation (K7025-05, Life Technologies, Paisly, UK). In-house designed barcode-adapters were ligated to both ends of the DNA fragments according to the standard New England Biolabs protocol using the NEBNext library prep master mix set (NEB, Ipswich, USA). Fragments of ~300bp were excised using the PerkinElmer labchipXT gel

system and DNA was extracted and amplified by PCR. Next, the PCR products of 4 independent samples were mixed in equimolar pools and used for exome enrichment using the Agilent SureSelect All exon V4 kit including approximately 51Mb of genomic sequences, according to the protocol of the manufacturer. PCR products were subjected to paired-end sequencing on the HiSeq2000. Image Files were processed using standard Illumina® base calling software and de-multiplexed using an in-house script.

All sequence reads were aligned to the human reference genome (build b37 released by the 1000 Genomes Project)[2] using Burrows-Wheeler Aligner.[3] Duplicate reads were marked by Picard (http://goo.gl/0sCehO). Using the Genome Analysis Toolkit (GATK)[4] reads mapping around insertions and deletions of the 1000 Genomes Project were re-aligned,[2] followed by a base quality score recalibration. During this process the quality of the data is assessed by FastQC (http://goo.gl/6TUqD), Picard, GATK Coverage and custom scripts. Single nucleotide variants (SNV) were called with GATK Unified Genotyper. Indels were called with GATK and Pindel[5] and those that were called with both programs are included in the final list. Variants were annotated using GATK and SNPEff.[6] This production pipeline was implemented using the MOLGENIS compute[7] platform for job generation, execution and monitoring (http://goo.gl/XLbc0F). In the output file we included amongst others Gene symbol, mutation position, amino acid change, all transcript IDs, SIFT score, Polyphen2 classification and genotype of the mutation.

A combination of different filtering steps was applied to further limit the list to functional mutations with adequate read depth. We removed all variants that (1) were present in the dbSNP release 135, (2) map in non-coding regions and (3) that are synonymous. For the remaining SNVs we next applied the following additional filtering steps to reduce the amount of putative false positives mutations from the final list: (1) a coverage of less than 10 reads, (2) a strand bias greater than or equal to zero; (3) an alternative allele read depth below 6; or an alternative allele fraction below 0.25. These filtering steps were implemented using a custom script.

***Comparison to the Broad data set***

The overlap in the captured regions between the Agilent WES kit and the kit used by the Broad Institute was 3,804,686bp (WES a total length of 51,189,318bp and for the Broad Institute a total length of 4,337,615bp). 20,978 genes are captured with the Agilent kit and 1,654 genes by the Broad institute, with an overlap of 1,584 genes (Supplementary Figure S2). For the 5 cell lines

that were analyzed by WES and the Broad institute, we compared the mutations focusing on the set of 1,584 genes. For inconsistencies we manually inspected reads of WES and Broad data. In case the mutant allele was present at low read numbers we considered the mutation to be consistent. In case the mutant allele was not present and the mutant allele position was not efficiently captured, the variant was called as inconclusive and excluded from the comparison for that specific cell line. In case sufficient reads were observed, but no mutant reads, the variant was called as truly inconclusive.

***Comparison to RNA-seq data***

RNA-seq libraries were constructed using protocols as previously described.[8] Sequencing was performed using a combination of Illumina Genome Analyzer II (KMH2, L428, DEV) and Illumina HiSeq 2000 instruments (SUPHD1, L540, L1236, L591) to produce paired-end reads of length 50-76. These libraries were then aligned to the UCSC hg19 using GSNAP[9] and multiple mapped reads were filtered using samtools. To examine if the somatic mutations were detectable in the RNA-seq data, we extracted the aligned reads using the Integrative Genomics Viewer (IGV) software. The mutation is considered to be verified if at least one read carried the mutant allele and the mutant allele was detected at a frequency >5% of the total reads.

***CNV analysis***

Pseudo probe data were generated with VarScan2 and Samtools as described.[10,11] Briefly, for each cell line the pseudo probe derived GC-normalized log2 copy number ratio was generated by comparing the read counts with the pileup of 3 normal PBMC samples. All alignments with a mapping quality greater than 40 in combination with a minimal segment size of 200 bp and a maximal segment size of 500 bp were used to calculate the log2 ratios. These log2 ratios were used for segment calling with the integrated DNA copy algorithm[12] in the Nexus 7 software package (Biodiscovery, USA). The deletion / duplication regions were filtered for a minimum of 20 probes.

***Linking WES data to copy number variations***

For each individual cell line, the list of names of the mutated genes was uploaded to Nexus as a custom track. This custom track was used to annotate each gene with the corresponding copy number separately for each cell line.

***Linking of WES data to differentially expressed genes***

Labeling and hybridization of RNA isolated from 5 cHL cell lines and from germinal B cells sorted from three independent donors was performed using one-color low input Quick Amp Labeling Kit, according to the manufacturer's protocol (Agilent, Santa Clara, USA) on a custom designed 8x60K sureprint G3

microarray that contained all protein coding gene probes present on the human gene expression v1 8x60k microarray (AMADID #028004, Agilent). Slides were scanned with GenePix 4000B (Agilent). Scanned images were used for Agilent Feature Extraction software version 10.7.3.1 and converted into Linear and Lowess normalized data. Using GeneSpring GX version 11.5.1 (Agilent), quantile normalization of the signals was performed. Next, 22,027 probes detected in at least 50% of the samples were included for analysis. An unpaired T-test in combination with Benjamini-Hochberg multiple testing correction was used to compare gene expression in cHL cell lines versus GC B cells. Venn diagrams (http://goo.gl/eV4Xw) were generated to check the overlap between genes mutated in at least 2 out of 5 classical HL cell lines and genes differentially expressed between cHL cell lines and GC B cells. Heatmaps were generated with Genesis software (v.1.7.6) by performing unsupervised complete linkage hierarchical clustering analysis (euclidean distance).

***Define gene ontologies of genes mutated in HL using DAVID***

Lists of mutated genes were uploaded and analyzed for gene annotations and gene ontology in the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (http://goo.gl/ERxui) using a previously reported protocol.[13] For genes with multiple gene ontologies or annotations, we preferentially annotated the gene ontology groups that were also used in the B cell lymphoma papers to which we compared our WES data.[14-19]

***Sanger sequencing***

Total RNA was isolated using standard laboratory protocols. cDNA was synthesized using 500ng input RNA using Superscript II according to the manufacturers protocol (Invitrogen, Carlsbad, USA). Genomic DNA or cDNA was amplified by PCR with AmpliTaq Gold® DNA Polymerase, PE Buffer II and $MgCl_2$ (Applied Biosystems, Foster City, USA) using primers designed using Primer Express (Applied Biosystems) (Supplemental Table S1). For part of the PCR products we linked an M13F or R tail, which was then used for sequencing. For the remaining PCR products the PCR primers were used for sequencing. PCR products were run on an agarose gel to check the amplification and purified using high pure PCR product purification kit (Roche, Mannheim, Germany) and sent for sequencing (LGC Genomics, Teddington, UK).

***qRT-PCR***

RNA isolation and cDNA synthesis were done as described above. For qRT-PCR we used 18S as a housekeeping gene for normalization and we show the relative expression values as $2^{-(\text{delta Cp})}$.

***FACS analysis***

Cells were incubated with mouse anti-HLA ABC (1:500, w6/32, MyBioSource, San Diego, CA, USA) and rabbit anti-human B2M (Dako, Denmark) antibodies for 30 min on ice. This was followed by a washing step with 1% BSA/PBS. Cells were incubated with secondary labeled antibodies (SouthernBiotech, Alabama, USA) for 30 min on ice, fixed with 2% paraformaldehyde PBS solution and analyzed on the BD FACSCalibur (BD Biosciences, New Jersey, USA).

## RESULTS AND DISCUSSION

After applying filtering steps and excluding known variants, we observed more than 1,000 mutated genes in SUPHD1, around 500 in L540, KMH2, L428 and L1236 and around 350 in DEV and L591 (Supplementary Tables S2 and S3). Overall, we identified 463 recurrently mutated genes in HL regardless of subtype (Figure 1A) and 373 recurrent mutations specifically in cHL (Supplementary Figure S1). Due to lack of constitutional DNA, we cannot rule out presence of personal variants in the list of somatic mutations. However, based on SNPEFF_IMPACT, PolyPhen2 and SIFT analyses we showed that approximately half of the mutations most likely have a damaging effect. It should be noticed that the cell lines have all been established from patients with end stage disease and as such might contain treatment-induced mutations in addition to possible cell line artifacts.

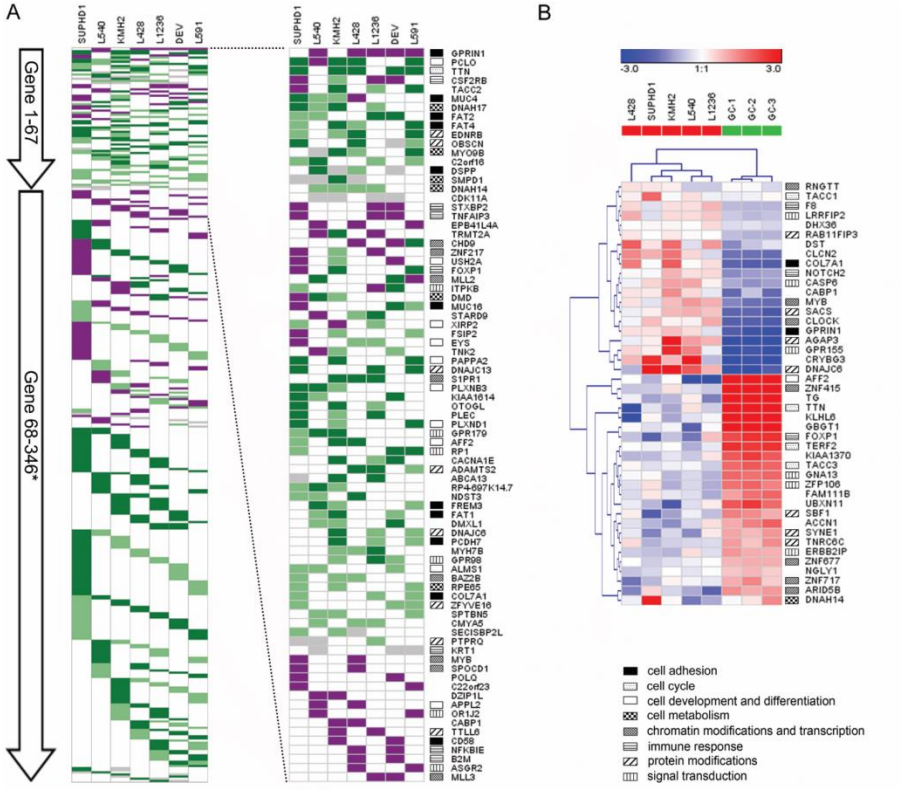



**Figure 1. Mutation patterns and gene ontology analysis. (A)** 463 genes were mutated in at least 2 out of 7 HL cell lines. Of these 346 are shown in the left panel based on a high impact mutation or a SIFT score ≤0.05 in at least one HL cell line. The 81 genes shown in the right panel includes 67 genes mutated in at least 3 HL cell lines and 14 genes showing high impact mutations in 2 HL cell lines. Results of the validation experiments by RNA-seq and Sanger sequencing are summarized in supplementary Table S7. Purple boxes indicate genes with mutations that have a SNPEFF_IMPACT score "high", including stop gain, frame shift and splice site mutations. Green and grey boxes indicate genes with SNPEFF_IMPACT "moderate", including missense mutations. Dark green boxes are mutations that probably are damaging based on SIFT score ≤0.05, light green boxes are mutations that are probably not damaging based on a SIFT score >0.05, grey boxes are mutations without SIFT score. (**B)** Heatmap of the gene expression levels of the 44 genes that were mutated specifically in cHL and differentially expressed in cHL vs GC B cells; 20 of the mutated genes were upregulated and 24 downregulated in cHL. The most common gene ontology classifications are indicated by boxes left of the gene names.

As a validation, we compared our WES data to the data from the Broad Institute that screened for the presence of mutations in a set of 1,654 genes in five of the cell lines (Supplementary Figure S2). Inconsistencies were checked and could be explained by low read numbers probably due to ineffective capturing. Truly inconsistent variants may represent differences that have been introduced during prolonged culturing and selective outgrowth of specific subclones. The overall consistency ranged from 72.1% in L428 to 98.5% in SUPHD1 for the 1584 genes analyzed in both datasets (Supplementary Table S3). Next, we made a comparison to previously reported gene mutations in individual HL cell lines (Supplementary Table S5). With the exception of the second deletion in *SOCS1* in L1236, we confirmed all mutations. By a combination of Sanger sequencing and RNA-seq data analysis we confirmed 128 out of 142 mutations shown in Figure 1B (Supplementary Table S7). Moreover, several genes previously reported to be mutated in primary cHL cases, were also mutated in the HL cell lines, e.g. *TNFAIP3*, *NFKBIE*, *CYLD* and *NFKBIA*.[1] More recently, *PTPN1* mutations were reported at a high frequency in cHL cases and cell lines.[20] Consistent with this, we observed *PTPN1* gene mutations in L428 and L1236 which were not called due to insufficient reads or strand bias. The high validation rate combined with the consistency with the Broad data and detection of mutations in genes previously reported to be mutated in HL cell lines and primary cases underscores the reliability and relevance of our WES analysis.

Two gene ontologies, cell adhesion (n=10) and cell development and differentiation (n=10), were most common among the genes that were either mutated in at least three cell lines (n=67) or showed high impact mutations in two HL cell lines (n=14). There are no previous studies on mutations in cell adhesion genes in HL previously. Especially the FAT and MUC gene families are of interest as multiple members were mutated. Genes of the FAT family encoding transmembrane proteins are frequently mutated across multiple human cancer types[21] and mutations in *MUC4* and *MUC6* were observed in colorectal cancer and childhood ALL.[22,23] As cellular interactions with the microenvironment are crucial to survival and outgrowth of HRS cells, it might be of interest to follow-up on these cell adhesion genes.

As loss of HLA expression is a common feature especially in EBV negative HL cases[24] and the mechanisms responsible for the HLA loss are largely unknown, we made an inventory of the mutations in HLA associated genes. We identified mutations in *B2M*, *HLA-A* and *CIITA* in one or more of the cell lines (Supplementary Table S6). Presence of *B2M* gene mutations was confirmed by Sanger sequencing at the DNA and mRNA level. Consistent with the *B2M*

mutations affecting the ATG start codon, we observed no or very low membrane B2M and HLA class I expression in L428 and DEV by flow cytometry (Figure 2). At the mRNA level, *B2M* levels were reduced in both cell lines as compared to L1236, whereas HLA-A, HLA-B and HLA-C levels were in the same range. The reduced *B2M* mRNA levels might be caused by a lower transcriptional activity or by a reduced stability of the mRNA transcript. Together our findings demonstrate that mutations affecting the ATG start codon of *B2M* in L428 and DEV contribute to the lack or very low membrane expression of membrane B2M and HLA class I. Our WES data revealed *CIITA* mutations in 2 of the 7 HL cell lines, i.e. L428 and L1236, albeit with insufficient reads. Both mutations were confirmed by Sanger sequencing and were also present in the Broad data set. In addition, it has been shown that this gene is inactivated due to a chromosomal translocation in KMH2.[25] *CIITA* encodes for the HLA class II transactivator and inactivation of this gene might explain downregulation of HLA class II expression frequently observed in primary cHL cases.

To further establish possible pathogenic relevance of the mutated genes we determined if these gene loci showed copy number changes and/or aberrant expression patterns. Copy number variation plots were generated of the HL cell lines using the total read counts (Supplementary Figure S3). A proportion of the genes mapped to chromosomal regions that showed copy number aberrations in HL (Supplementary Table S7), but there was no significant enrichment of mutated genes in these regions. Gene expression data available for five cHL cell lines, i.e. L428, L540, L1236, SUPHD1 and KMH2, and three sorted GC B cell samples, were used to establish a possible link between mutation status and differential expression. In total 2,362 genes were significantly differentially expressed in cHL vs GC B cells after multiple testing correction. Of the 373 genes mutated specifically in at least two of the five cHL cell lines, 20 showed a significant up- and 24 a significant downregulation in cHL compared to GC B cells (Figure 1B and Supplementary Table S7). There was no significant enrichment of mutated genes in the list of differentially expressed genes. The two most common gene ontologies in this group of 44 differentially expressed genes were protein modifications (n=7) and chromatin modifications and transcription (n=7).

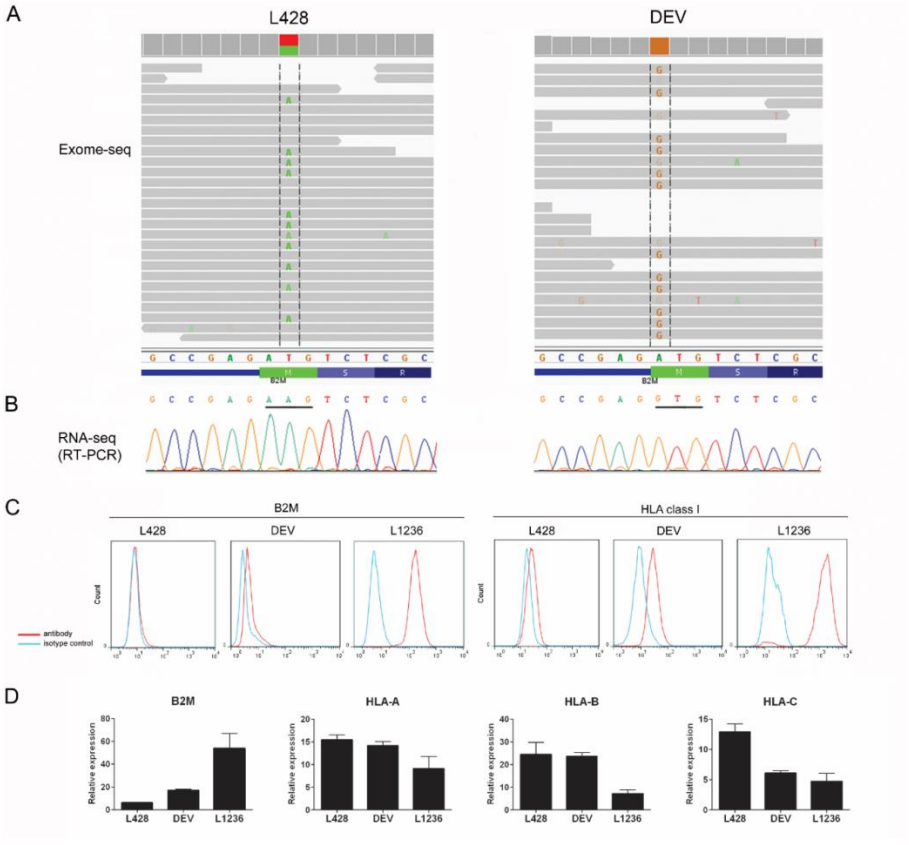



**Figure 2. Mutations affecting the ATG start codon of *B2M* in L428 and DEV, resulting in loss of B2M protein expression. (A)** Translation start-site loss for *B2M* in L428 and DEV identified in WES data. **(B)** Confirmation of the mutations by Sanger sequencing at the RNA level. No wild type allele is observed by Sanger sequencing. The altered start codon is underlined. **(C)** FACS staining of *B2M* and HLA class I in L428 and DEV. As a control, L1236 cells showed normal positive staining of B2M and HLA class I. **(D)** Quantitative RT-PCR analysis of B2M, HLA-A, HLA-B and HLA-C in the same three HL cell lines.

Although the number of WES or whole genome sequencing (WGS) studies in NHL is still quite limited, we explored the potential overlap of the mutation pattern observed in HL with that of other germinal center B cell derived NHL subtype[8,14-19] (Supplementary Table S8). Overall, we observed a large overlap with genes mutated in DLBCL and FL, whereas the overlap with BL was limited (Supplementary Figure S4). 88 genes mutated in HL were also mutated in at least one of the other lymphoma subtypes. The most common gene ontology identified within the list of 88 overlapping genes is the chromatin modification

and transcription gene ontology with 13 genes, e.g. *CREBBP*, *MLL2*, *MLL3*, *BCL6*, *HIST1H1E* and *EP300*. Mutations of histone-modifying genes are also common in DLBCL and FL. The common origin of HRS cells and the tumor cells of DLBCL and FL is consistent with the marked overlap in the mutational spectrum. For BL, the overlap is limited, and further genomic studies are awaited to clearly establish the extent of overlap. Despite the significant overlap, the majority of the genes that were mutated in HL remain unique. This indicates that the genetic events occurring during the malignant transformation in HL are distinct from those in other GC B cell derived non-Hodgkin lymphomas.

In conclusion, we determined the mutational landscape of HL cell lines by WES. Several of the mutated genes are in common with other germinal center B-cell derived lymphoma subtypes. We observed a high number of consistently mutated genes, with part of them mapping to regions with copy number gain or loss and part of them showing deregulated expression. This supports their potential relevance in HL and will help to select candidate genes for further analysis.

**Acknowledgements**
We acknowledge the Genome facility of the Genetics Department (Pieter van der Vlies and others). This work was supported by grants from the Abel Tasman Talent Program (ATTP), University Medical Center Groningen and the Dutch Cancer Society (KWF: RUG 2009-4313).

## Supplementary Information

**Supplementary Table S1.** Primers for PCR validation.

| Gene (material) | Forward / reverse | Primer Sequence (5'-3') M13-tail for sequencing is shown underlined |
|---|---|---|
| B2M_(qRT-PCR) | forward | GAAAAAGTGGAGCATTCAGACTTG |
| | reverse | ATGATGCTGCTTACATGTCTCGAT |
| HLA-A (qRT-PCR) | forward | GATGTGTCCCTCACAGCTTGTAAA |
| | reverse | CAGGAACAACTCTTGCCTCTCA |
| HLA-B (qRT-PCR) | forward | CTAGCAGTTGTGGTCATCGGAG |
| | reverse | TCAAGCTGTGAGAGACACATCAGAG |
| HLA-C (qRT-PCR) | forward | GAGGAAGAGCTCAGGTGGAAA |
| | reverse | AGCCCTGGGCACTGTTG |
| 18S (qRT-PCR) | forward | CGGCTACCACATCCAAGGA |
| | reverse | CCAATTACAGGGCCTCGAAA |
| B2M exon 1_DEV & L428 (RNA) | forward | GTAAAACGACGGCCAGAATATAAGTGGAGGCGTCGCG |
| | reverse | GGAAACAGCTATGACCATGCCCAGACACATAGCAATTCAGG |
| B2M exon 1_DEV & L428 (DNA) | forward | GTAAAACGACGGCCAGCTAACCTGGCACTGCGTCG |
| | reverse | GGAAACAGCTATGACCATGAGCGAGAGAGCACAGCGAG |
| CIITA exon 1_L428 & L1236 (RNA) | forward | GTAAAACGACGGCCAGTGATGAGGCTGTGTGCTTCTG |
| | reverse | GGAAACAGCTATGACCATGCACAGGGGGTCAGCATCG |
| CIITA exon 4-5_L428 & L1236 (DNA) | forward | GTAAAACGACGGCCAGCCAAGGCCTGGCACACAG |
| | reverse | GGAAACAGCTATGACCATGCCACAGCAGGCTTTGGAGTC |
| CIITA exon 7_L428 & L1236 (DNA) | forward | GTAAAACGACGGCCAGCCAGGGCTGAGAAGATGACAAG |
| | reverse | GGAAACAGCTATGACCATGGGGCAGAGGTGAGTGACACTG |
| CIITA exon 18_ L428 & L1236 (DNA) | forward | GTAAAACGACGGCCAGGAAGTCTCTCCCTGGTGTCCTG |
| | reverse | GGAAACAGCTATGACCATGGATTGGTGAAGGGAAGAGCATG |
| C2orf16_L1236 | forward | AGGAACGGTCCAGTGCAGAT |
| | reverse | AGGAACTGCGTTGGCTTCTC |
| C2orf16_L540 | forward | GCAAGGACACAGCTTCAAGT |
| | reverse | GAAGATGTGGCCTTGGAAGT |
| C2orf16_L591 | forward | CTACATCAGGCGAGGACAAC |
| | reverse | TGCTCCAATCCTCTTCCTTC |
| C2orf16_SUPHD1 | forward | GGACATCGAGTTCCTGAATC |
| | reverse | CTAGTGGCTCTGGACTAATC |
| DNAH17_L1236 | forward | AGGCCAACGAGGTCAGCATC |
| | reverse | GCATGTGCTCGCACACGTAT |
| DNAH17_L540 | forward | TGACGCTGAGCTAGACAATG |
| | reverse | CACAGCCTCTTCCTGTTCCT |
| DNAH17_SUPHD1 | forward | CTGGAGAAGCCGCTGGAGAA |
| | reverse | TCCATGTGCTGCCGGATGAG |
| DSPP_L540 | forward | GGACACAGCAATACAGGTAG |
| | reverse | TTGTTACCATTGCCATTACT |
| EDNRB_KMH2 | forward | CTGGTTGCGCTGGTTCTTGC |
| | reverse | AGGACACAACCGTGTTGATG |
| EDNRB_L428 | forward | GTCACTTCGGTTCCACTTCA |
| | reverse | CTCTCAACAGGACCTCAGAT |
| EDNRB_L591 | forward | GTCACTTCGGTTCCACTTCA |
| | reverse | CTCTCAACAGGACCTCAGAT |
| EDNRB_SUPHD1 | forward | CTGGAGAGGAGGAAGAGTTG |
| | reverse | GCACCAGGCAGTTCACAGTC |
| FAT2_L1236 | forward | ACAGGCTGAAGGTTCCAGAG |
| | reverse | TTCACATCCAGGACGATCAC |
| FAT2_L540 | forward | CAGATCAAGGCAACAGACAG |
| | reverse | CGACCTTGAGAAGAGTCAGA |

| | | |
|---|---|---|
| FAT4_KMH2 | forward | ACAGATGCCAAGAGGCAACC |
| | reverse | GTCAGAACGGAACCAATCAC |
| FAT4_L428 | forward | TAGCCTGATCCTGTGTAACC |
| | reverse | GACTGTGCCTCTGGATTGAA |
| FAT4_L591 | forward | CCATGGCATCACATGGTTCT |
| | reverse | GCTTCAGTCTGGCTCCATAA |
| FAT4_SUPHD1 | forward | ACTCCTGACACTGCCTTATC |
| | reverse | ACGGCAGGCGGTCTAATCAA |
| GPRIN1_DEV | forward | GATCCTGTGGCTCCTGGAAG |
| | reverse | GTGGACGCAGAATCCATGCT |
| GPRIN1_L428 | forward | GATCCTGTGGCTCCTGGAAG |
| | reverse | GTGGACGCAGAATCCATGCT |
| MUC4_KMH2 | forward | GTGGAGACCACCAGAGTATC |
| | reverse | GAGGAAGGCCATGTTGTTGT |
| MUC4_L428 | forward | CTCAGCATCCACAGGTCACG |
| | reverse | TAGGCTGAATTCCGCCAAGG |
| MUC4_L540 | forward | GGAGACAGCTCCTCCAGATG |
| | reverse | AGTCCAGTGGTTCTGAGAAG |
| MUC4_SUPHD1 | forward | TCCTTCCAGCCTGGATGTTC |
| | reverse | TCCTTCTCCTCGGCCTCAGT |
| MYO9B_L591 | forward | GTCATCAAGGACGCCATTGC |
| | reverse | CGGTGCTTGAGGTTCTTCAG |
| OBSCN_DEV | forward | GTGAGGTCATCTGGCACAAG |
| | reverse | CCTGTGGAGCCTCAGTCTCT |
| OBSCN_L591 | forward | CAAGAGGCTGAAGACAATGG |
| | reverse | ACAATGAGTGAGGCCTTGGA |
| PCLO_KMH2 | forward | TCCGATGGAATATCAAGCTC |
| | reverse | AACTTCATGGGGTTGTGTTT |
| PCLO_L428 | forward | GCACCTGTTACCACTACATC |
| | reverse | GATGGGACAGAAGGTTCTTT |
| PCLO_L540 | forward | ATCACGGACTCGAGGCTATG |
| | reverse | TGCAAAAGACACACCAAGAA |
| PCLO_L591 | forward | CTGTGGCAACAATGTCCTTC |
| | reverse | GCTTAATGCCGCTGTTCCTT |
| PCLO_SUPHD1 | forward | ACCAGATGGTAGAGCTAGTG |
| | reverse | CGTTCGGCCTCCAACTCCTT |
| SMPD1_L428 | forward | GTGTAGGAAGCGCGACAATG |
| | reverse | GCGATGTAACCTGGCAGGAT |
| SMPD1_L540 | forward | GTGTAGGAAGCGCGACAATG |
| | reverse | CGATGTAACCTGGCAGGATG |
| SMPD1_SUPHD1 | forward | GTGTAGGAAGCGCGACAATG |
| | reverse | CGATGTAACCTGGCAGGATG |
| TACC2_L591 | forward | GCAGTGTGGCTCTGTCCTTG |
| | reverse | TGAGTCTGAGGACCATCTTC |
| TACC2_SUPHD1 | forward | TGATGGTGCTTCTTCCTCAG |
| | reverse | CTTCCGTCTTGGCAGATTCC |
| TTN_DEV | forward | CCAAGTGTGCTGGTATAGAG |
| | reverse | CGGTTGAGCACCAGTAACAT |
| TTN_L428 | forward | GTGCAAGGTTGGCTGGAGAC |
| | reverse | CAGCTGTGATGGCTCTAACT |
| TTN_L591 | forward | CCGCAGTTGAAGCACTTGAC |
| | reverse | AATCTGCAGCTGTAGGTTGG |

**Supplementary Table S2.** Overview of coverage and number of mutations of our WES data.

| Cell line | SUPHD1 | L540 | KMH2 | L428 | L1236 | DEV | L591 |
|---|---|---|---|---|---|---|---|
| Total reads (x10$^6$) | 87 | 175 | 83 | 115 | 105 | 121 | 112 |
| Mean coverage[a] | 51 | 107 | 50 | 66 | 79 | 72 | 58 |
| 2x coverage[b] | 98.4% | 98.6% | 98.2% | 98.4% | 98.9% | 98.9% | 98.5% |
| 10x coverage[b] | 90.8% | 96.5% | 89.6% | 94.0% | 95.1% | 95.6% | 93.6% |
| 20x coverage[b] | 76.1% | 92.1% | 74.0% | 84.2% | 86.5% | 87.5% | 82.4% |
| 30x coverage[b] | 61.0% | 85.8% | 58.8% | 72.4% | 76.1% | 76.9% | 68.9% |
| # mutations[c] | 1,192 | 521 | 524 | 501 | 491 | 377 | 357 |
| # mutated genes | 1,104 | 502 | 489 | 474 | 468 | 363 | 340 |
| SNPEFF_IMPACT | | | | | | | |
| High | 195 | 81 | 70 | 75 | 68 | 62 | 42 |
| Moderate | 997 | 440 | 454 | 426 | 423 | 315 | 115 |
| SIFT scores | | | | | | | |
| ≤0.05 | 433 | 173 | 174 | 179 | 172 | 142 | 109 |
| >0.05 | 562 | 250 | 268 | 227 | 233 | 174 | 196 |
| not available | 197 | 98 | 82 | 95 | 86 | 61 | 52 |

[a]of all baits; [b]fraction of target covered by; [c]a complete list of all individual mutations, including nucleotide position, amino acid change, SNPEFF, PolyPhen2 and SIFT scores per cell line is given in Supplementary Tables S3.

**Supplementary Table S3.** All somatic mutations of our WES data in 7 HL cell lines

Available online (http://www.nature.com/leu)

**Supplementary Table S4.** Summary of the comparison between WES data and reported data from the Broad Institute.

| Cell line | Consistently called by pipeline | Only called in WES | | | Only called in Broad | | | Overall consistency** |
|---|---|---|---|---|---|---|---|---|
| | | After manual inspection in Broad* | | | After manual inspection in WES* | | | |
| | | Consistent | Inconclusive | Inconsistent | Consistent | Inconclusive | Inconsistent | |
| SUPHD1 | 113 | 8 | 4 | 2 | 12 | 27 | 0 | 133/135 (98.5%) |
| L540 | 49 | 0 | 3 | 1 | 17 | 17 | 0 | 66/67 (98.5%) |
| KMH2 | 47 | 0 | 8 | 1 | 6 | 20 | 0 | 53/54 (98.1%) |
| L428 | 37 | 5 | 3 | 13 | 7 | 12 | 6 | 49/68 (72.1%) |
| L1236 | 38 | 0 | 6 | 6 | 7 | 11 | 4 | 45/55 (81.8%) |

* For inconsistencies we manually inspected reads of WES and Broad data. In case, the mutant allele was present at low read numbers we considered the mutation to be consistent. In case, the mutant allele was not present and the mutant allele position was not efficiently captured, the variant was excluded from the comparison for that specific cell line, i.e. inconclusive. In case the mutant position was efficiently captured but the mutant allele was not detected we considered it as inconsistent.

** The overall consistency was calculated using the formula: (consistently called + consistent after inspection) / (consistently called + consistent after inspection + inconsistent) x100%.

**Supplementary Table S5**. Consistency of our whole exome seq data and data from the Broad Institute with current literature on single genes.

| cell line | Gene | Reported mutation | WES | Broad |
|---|---|---|---|---|
| **KMH2** | CYLD[26] | g54506delT, g.54508G>A | yes | WT |
| **KMH2** | NFKBIA[27-29] | del part exon 3 / intron 3 del 509-641 | yes | WT |
| **KMH2** | TNFAIP3[30,31] | large deletion intron 2 - exon 6 | no reads* | WT |
| **L1236** | FAS[32] | splice donor intron 7 | yes | yes |
| **L1236** | SOCS1[33] | Deletions 12 nt and 28 nt | 28nt del. | unknown |
| **L1236** | STAT6[34] | N417Y | yes | yes |
| **L1236** | TNFAIP3[30,31] | G491A results W142STOP | yes | yes |
| **L1236** | TP53[35] | Loss of exon 10-11 | no reads* | WT |
| **L428** | NFKBIA[27-29] | C-T exon 5 893C>U | yes | yes |
| **L428** | NFKBIE[36] | hemizygous frame shift | yes | unknown |
| **L428** | SOCS1[33] | 11 nt and 15 nt deletions | yes | WT |
| **L428** | TP53[35] | Loss 33bp exon 4 | yes | WT |
| **SUPHD1** | PTPN2[27] | hemizygous missense mutation in exon 7 | yes | WT |

WT, wild type sequence; * lack of reads for the exons previously reported to be deleted is consistent with presence of the large deletions.

**Supplementary Table S6.** Overview of the mutated HLA and associated genes.

| Cell lines | Gene | Mutation | Effect on protein |
|---|---|---|---|
| DEV | β2M | M1V | loss of translation start site |
| L428 | β2M | M1K | loss of translation start site |
| SUPHD1 | HLA-A | CACG insertion in exon 1 | frame shift |
| L428 | CIITA | S11P* | missense |
| L1236 | CIITA | R2C* | missense |

*The two mutations in CIITA were not called in our WES pipeline due to insufficient reads, but were validated by Sanger sequencing and reported in the Broad data set.

**Supplementary Table S7.** Overview of mutations, differential expression, copy number variations and validation results per cell line.

| gene symbol | mutated | mutated | diff. expr. | SUPHD1 | | L540 | | KMH2 | | L428 | | L1236 | | DEV | | L591 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | in >3 | in >2 cHL | in cHL | WES | CNV | WES | CNV | WES | CNV | WES | CNV | WES | CNV | WES | CNV | WES | CNV |
| GPRIN1 | yes | yes | yes | WT | - | Mut. | Gain (hc) | WT | - | Mut. | Gain (hc) | Mut. | Gain | Mut. | - | Mut. | - |
| PCLO | yes | yes | no | Mut. | - | Mut. | Gain (hc) | Mut. | Gain | Mut. | Gain | WT | - | WT | - | Mut. | Gain |
| TTN | yes | yes | yes | Mut. | - | WT | - | Mut. | Gain | Mut. | Gain | WT | - | Mut. | - | Mut. | Gain |
| CSF2RB | yes | yes | no | Mut. | - | WT | - | Mut. | - | WT | - | Mut. | - | Mut. | - | WT | - |
| TACC2 | yes | yes | no | Mut. | - | WT | - | Mut. | - | WT | - | Mut. | - | WT | - | Mut. | - |
| MUC4 | yes | yes | no | Mut. | - | Mut. | Gain | Mut. | Gain (hc) | Mut. | Gain | WT | - | WT | - | WT | - |
| DNAH17 | yes | yes | no | Mut. | - | Mut. | Gain | Mut. | - | WT | - | Mut. | - | WT | - | WT | - |
| FAT2 | yes | yes | no | Mut. | - | Mut. | Gain | WT | - | WT | - | Mut. | - | Mut. | - | WT | - |
| FAT4 | yes | yes | no | Mut. | Loss | WT | - | Mut. | - | Mut. | Gain | WT | - | WT | - | Mut. | - |
| EDNRB | yes | yes | no | Mut. | - | WT | - | Mut. | - | Mut. | Loss | WT | - | WT | - | Mut. | - |
| OBSCN | yes | yes | no | Mut. | Gain | WT | - | WT | - | Mut. | Gain | WT | - | Mut. | - | Mut. | - |
| MYO9B | yes | yes | no | WT | - | Mut. | Gain (hc) | Mut. | Gain (hc) | Mut. | Gain | WT | - | WT | - | Mut. | - |
| C2orf16 | yes | yes | no | Mut. | - | Mut. | Gain (hc) | WT | - | WT | - | Mut. | Gain (hc) | WT | - | Mut. | Gain |
| DSPP | yes | yes | no | WT | - | Mut. | Gain | Mut. | Gain (hc) | Mut. | Gain | WT | - | Mut. | - | WT | - |
| SMPD1 | yes | yes | no | Mut. | - | Mut. | Gain | Mut. | Gain | Mut. | Gain | WT | - | WT | - | WT | - |
| DNAH14 | yes | yes | yes | WT | - | Mut. | Gain | Mut. | Gain (hc) | Mut. | Gain | Mut. | Gain | WT | - | WT | - |
| CDK11A | yes | yes | no | WT | - | Mut. | Gain | Mut. | Gain (hc) | WT | - | Mut. | Gain (hc) | Mut. | - | WT | - |
| STXBP2 | yes | yes | no | Mut. | - | WT | - | WT | - | WT | - | Mut. | - | Mut. | - | WT | - |
| TNFAIP3 | yes | yes | no | Mut. | - | WT | - | WT | - | WT | - | Mut. | - | Mut. | - | WT | - |
| EPB41L4A | yes | yes | no | WT | - | Mut. | Gain | WT | - | Mut. | Gain | WT | - | WT | - | Mut. | - |
| TRMT2A | yes | yes | no | WT | - | Mut. | Gain | Mut. | Gain | WT | - | Mut. | - | WT | - | WT | - |
| CHD9 | yes | yes | no | WT | - | WT | - | WT | - | Mut. | Gain | WT | - | Mut. | - | Mut. | - |
| ZNF217 | yes | yes | no | Mut. | - | WT | - | Mut. | - | WT | - | Mut. | Gain | WT | - | WT | - |
| USH2A | yes | yes | no | Mut. | Gain (hc) | WT | - | Mut. | Gain (hc) | WT | - | WT | - | Mut. | - | WT | - |
| FOXP1 | yes | yes | yes | Mut. | - | WT | - | Mut. | Gain | WT | - | WT | - | WT | - | Mut. | - |
| MLL2 | yes | yes | no | WT | - | Mut. | Gain | WT | - | WT | - | WT | - | Mut. | Gain | Mut. | - |
| ITPKB | yes | yes | no | Mut. | Gain (hc) | WT | - | WT | - | WT | - | Mut. | Gain | Mut. | - | WT | - |
| DMD | yes | yes | no | Mut. | Loss | Mut. | Gain | Mut. | Gain | WT | - | WT | - | WT | - | WT | - |
| MUC16 | yes | yes | no | Mut. | - | WT | - | WT | - | WT | - | WT | - | Mut. | - | Mut. | - |
| STARD9 | yes | yes | no | WT | - | Mut. | Gain | WT | - | Mut. | Gain | Mut. | Gain | WT | - | WT | - |
| XIRP2 | yes | yes | no | Mut. | - | WT | - | Mut. | Gain | WT | - | Mut. | Gain | WT | - | WT | - |
| FSIP2 | yes | yes | no | Mut. | - | WT | - | Mut. | - | WT | - | WT | - | WT | - | Mut. | Gain |
| EYS | yes | yes | no | Mut. | - | WT | - | WT | - | Mut. | Gain | Mut. | Gain | WT | - | WT | - |
| TNK2 | yes | yes | no | WT | - | Mut. | Gain | Mut. | Gain (hc) | WT | - | WT | - | Mut. | - | WT | - |

| gene symbol | mutated | mutated | diff. expr. | SUPHD1 | | L540 | | KMH2 | | L428 | | L1236 | | DEV | | L591 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | in >3 | in >2 cHL | in cHL | WES | CNV | WES | CNV | WES | CNV | WES | CNV | WES | CNV | WES | CNV | WES | CNV |
| PAPPA2 | yes | yes | no | Mut. | Gain | WT | - | Mut. | Gain | WT | - | WT | - | WT | - | Mut. | - |
| DNAJC13 | yes | yes | no | Mut. | - | WT | - | WT | - | WT | - | Mut. | Gain (hc) | WT | - | Mut. | - |
| S1PR1 | yes | yes | no | WT | - | WT | - | Mut. | Gain | Mut. | Gain (hc) | Mut. | - | WT | - | WT | - |
| PLXNB3 | yes | yes | no | Mut. | Loss | Mut. | Gain | Mut. | Gain | WT | - | WT | - | WT | - | WT | - |
| KIAA1614 | yes | yes | no | Mut. | Gain | Mut. | Gain | WT | - | WT | - | WT | - | Mut. | - | WT | - |
| OTOGL | yes | yes | no | Mut. | - | WT | - | Mut. | Gain | WT | - | Mut. | - | WT | - | WT | - |
| PLEC | yes | yes | no | Mut. | Gain | WT | - | Mut. | Gain | WT | - | Mut. | Gain (hc) | WT | - | WT | - |
| PLXND1 | yes | yes | no | Mut. | - | WT | - | Mut. | Gain | WT | - | WT | - | WT | - | Mut. | - |
| GPR179 | yes | yes | no | Mut. | - | Mut. | Gain (hc) | Mut. | - | WT | - | WT | - | WT | - | WT | - |
| AFF2 | yes | yes | yes | Mut. | Loss | WT | - | Mut. | Gain | Mut. | Gain (hc) | WT | - | WT | - | WT | - |
| RP1 | yes | yes | no | Mut. | Gain | WT | - | WT | - | WT | - | WT | - | Mut. | - | Mut. | - |
| CACNA1E | yes | yes | no | WT | - | WT | - | Mut. | - | WT | - | Mut. | Gain | Mut. | - | WT | - |
| ADAMTS2 | yes | yes | no | WT | - | WT | - | WT | - | Mut. | Gain (hc) | Mut. | - | WT | - | Mut. | - |
| ABCA13 | yes | yes | no | Mut. | - | WT | - | Mut. | - | WT | - | Mut. | Gain | WT | - | WT | - |
| HELZ2 | yes | yes | no | Mut. | - | Mut. | Gain | Mut. | Gain (hc) | WT | - | WT | - | WT | - | WT | - |
| NDST3 | yes | yes | no | Mut. | Loss | Mut. | Gain | WT | - | Mut. | Gain | WT | - | WT | - | WT | - |
| FREM3 | yes | yes | no | WT | - | Mut. | Gain | Mut. | Gain | WT | - | WT | - | WT | - | Mut. | Loss |
| FAT1 | yes | yes | no | WT | - | Mut. | Gain | Mut. | Loss | WT | - | WT | - | Mut. | - | WT | - |
| DMXL1 | yes | yes | no | WT | - | Mut. | Gain | Mut. | - | WT | - | WT | - | Mut. | - | WT | - |
| DNAJC6 | yes | yes | yes | WT | - | WT | - | Mut. | Gain | WT | - | Mut. | Gain | WT | - | Mut. | - |
| PCDH7 | yes | yes | no | WT | - | WT | - | Mut. | Gain | WT | - | Mut. | - | WT | - | Mut. | Loss |
| MYH7B | yes | yes | no | WT | - | WT | - | Mut. | - | Mut. | Gain | Mut. | Gain (hc) | WT | - | WT | - |
| GPR98 | yes | yes | no | WT | - | WT | - | Mut. | Loss | WT | - | Mut. | - | Mut. | - | WT | - |
| ALMS1 | yes | yes | no | Mut. | - | Mut. | Gain | WT | - | WT | - | WT | - | Mut. | - | WT | - |
| BAZ2B | yes | yes | no | Mut. | - | WT | - | Mut. | Gain | Mut. | Gain | WT | - | WT | - | WT | - |
| RPE65 | yes | yes | no | Mut. | Gain | WT | - | Mut. | Gain | WT | - | WT | - | WT | - | Mut. | - |
| COL7A1 | yes | yes | yes | Mut. | - | WT | - | WT | - | WT | - | Mut. | Gain | WT | - | Mut. | - |
| ZFYVE16 | yes | yes | no | Mut. | - | WT | - | WT | - | WT | - | WT | - | Mut. | - | Mut. | - |
| SPTBN5 | yes | yes | no | WT | - | Mut. | Gain | Mut. | Gain | WT | - | WT | - | WT | - | Mut. | - |
| CMYA5 | yes | yes | no | WT | - | Mut. | Gain | WT | - | Mut. | - | Mut. | - | WT | - | WT | - |
| SEC1SBP2L | yes | yes | no | Mut. | - | WT | - | Mut. | Gain | Mut. | Gain | WT | - | WT | - | WT | - |
| PTPRQ | yes | yes | no | Mut. | - | Mut. | Gain | WT | - | WT | - | Mut. | - | WT | - | WT | - |
| KRT1 | yes | yes | no | WT | - | Mut. | Gain | WT | - | WT | - | WT | - | Mut. | Gain | Mut. | - |
| MYB | no | yes | yes | Mut. | - | WT | - | WT | - | Mut. | Gain | WT | - | WT | - | WT | - |
| SPOCD1 | no | yes | no | Mut. | Loss | WT | - | WT | - | Mut. | Gain | WT | - | WT | - | WT | - |
| POLQ | no | no | no | Mut. | - | WT | - | WT | - | WT | - | WT | - | Mut. | - | WT | - |

| gene symbol | mutated | mutated | diff. expr. | SUPHD1 | | L540 | | KMH2 | | L428 | | L1236 | | DEV | | L591 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | in >3 | in >2 cHL | in cHL | WES | CNV | WES | CNV | WES | CNV | WES | CNV | WES | CNV | WES | CNV | WES | CNV |
| C22orf23 | no | yes | no | Mut. | - | WT | - | WT | - | WT | - | WT | - | WT | - | Mut. | - |
| DZIP1L | no | yes | no | WT | - | Mut. | Gain | Mut. | Gain | WT | Gain | WT | Gain (hc) | WT | - | WT | - |
| APPL2 | no | yes | no | WT | - | Mut. | Gain | WT | Gain (hc) | Mut. | - | WT | - | WT | - | WT | - |
| OR1J2 | no | yes | no | WT | - | Mut. | - | WT | - | WT | - | WT | - | WT | - | Mut. | - |
| CABP1 | no | yes | yes | WT | - | WT | - | Mut. | Gain (hc) | Mut. | Loss | WT | - | WT | - | WT | - |
| TTLL6 | no | yes | no | WT | - | WT | - | Mut. | - | WT | - | Mut. | Gain | WT | - | WT | - |
| CD58 | no | no | no | WT | - | WT | - | Mut. | - | WT | Gain (hc) | WT | - | Mut. | - | WT | - |
| NFKBIE | no | no | no | WT | - | WT | - | WT | - | Mut. | Gain | WT | - | Mut. | - | WT | - |
| B2M | no | no | no | WT | - | WT | Gain | WT | Gain | Mut. | Gain | WT | Gain | Mut. | - | WT | - |
| ASGR2 | no | yes | no | WT | - | WT | - | WT | - | Mut. | Gain (hc) | WT | - | WT | - | Mut. | Loss |
| MLL3 | no | no | no | WT | - | WT | Gain (hc) | WT | Gain (hc) | WT | Gain | Mut. | - | Mut. | - | WT | Loss |

WES: whole exome sequencing result

CNV: copy number variations determined based on number of reads

Gain (hc): high copy number gain

WT: wild type

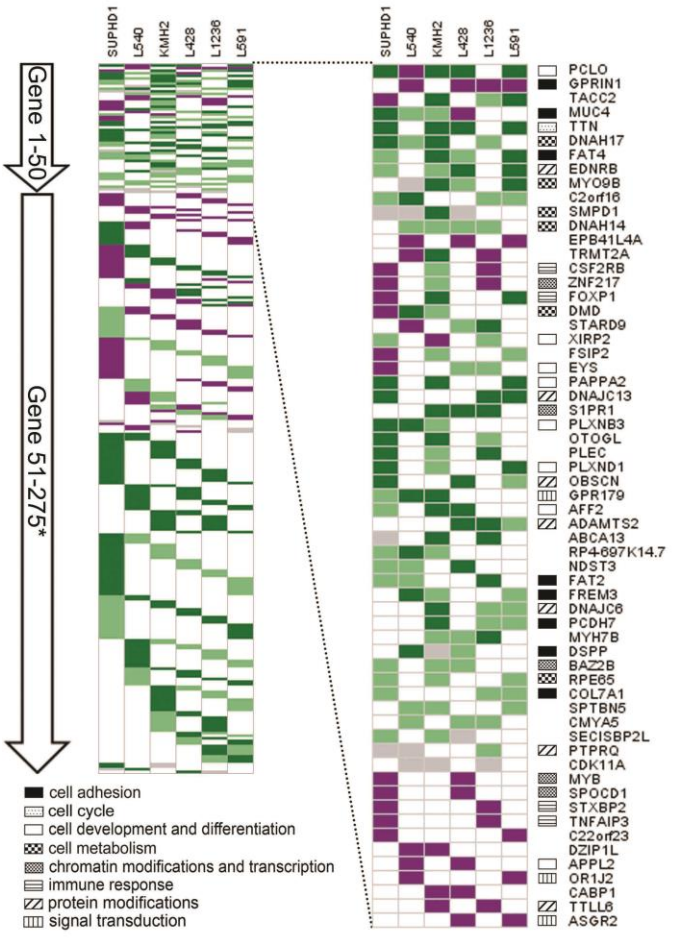
Mut: mutation

| | |
|---|---|
| Mut. | mutation confirmed with RNA-seq |
| Mut. | mutation confirmed with DNA Sanger sequencing |
| Mut. | mutation not confrimed with RNA-seq |
| Mut. | mutation not confirmed by both RNA-seq and DNA Sanger sequencing |
| Mut. | mutation for which no validation data are available due to insufficient reads by RNA-seq (<10) |

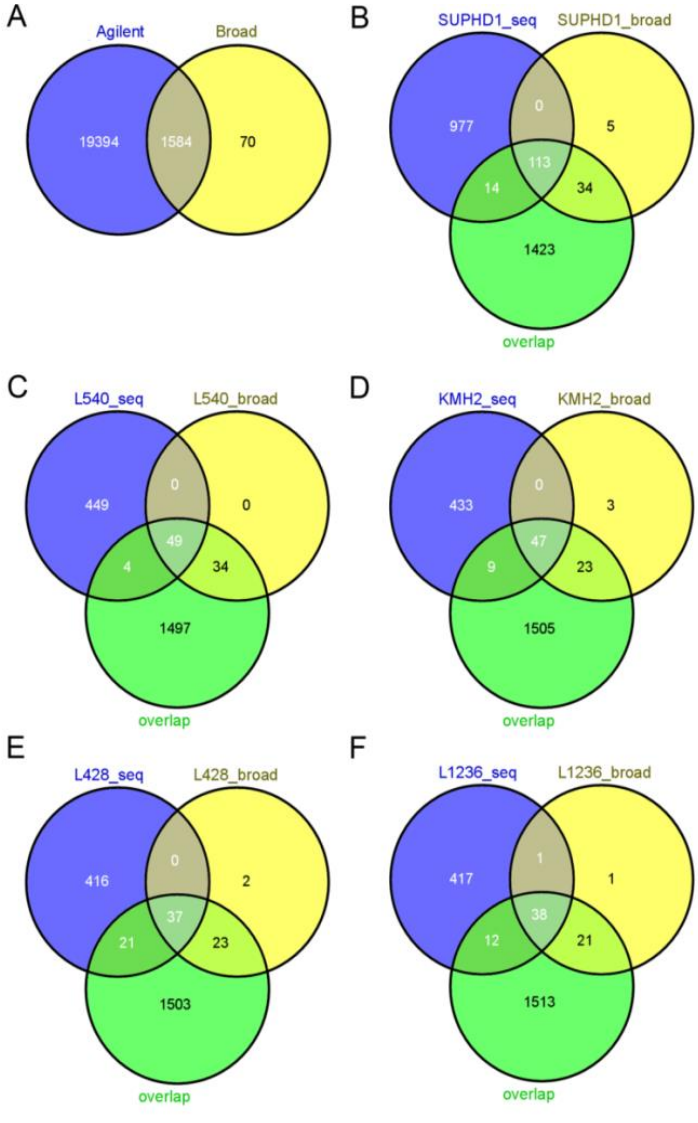
**Supplementary Table S8.** Selection of genes mutated in HL and NHL for identification of the overlap

| NHL subtype | Number of samples | No of mutated genes | criterion |
|---|---|---|---|
| DLBCL[14] | 117 cases and 10 cell lines | 109 | na |
| DLBCL[8] | 40 cases, 13 cell lines | 569 | na |
| DLBCL[15] | 6 cases | 23 | na |
| DLBCL[16] | 49 cases | 701 | na |
| DLBCL[17] | 94=73 cases, 21 DLBCL cell lines | 322 | na |
| DLBCL | Combined 285 cases | 499 | 2/285 |
| FL[18] | 8 cases | 569 | 1/8 |
| BL[19] | 4 cases | 102 | 1/4 |
| HL | 7 cell lines | 463 | 2/7 |

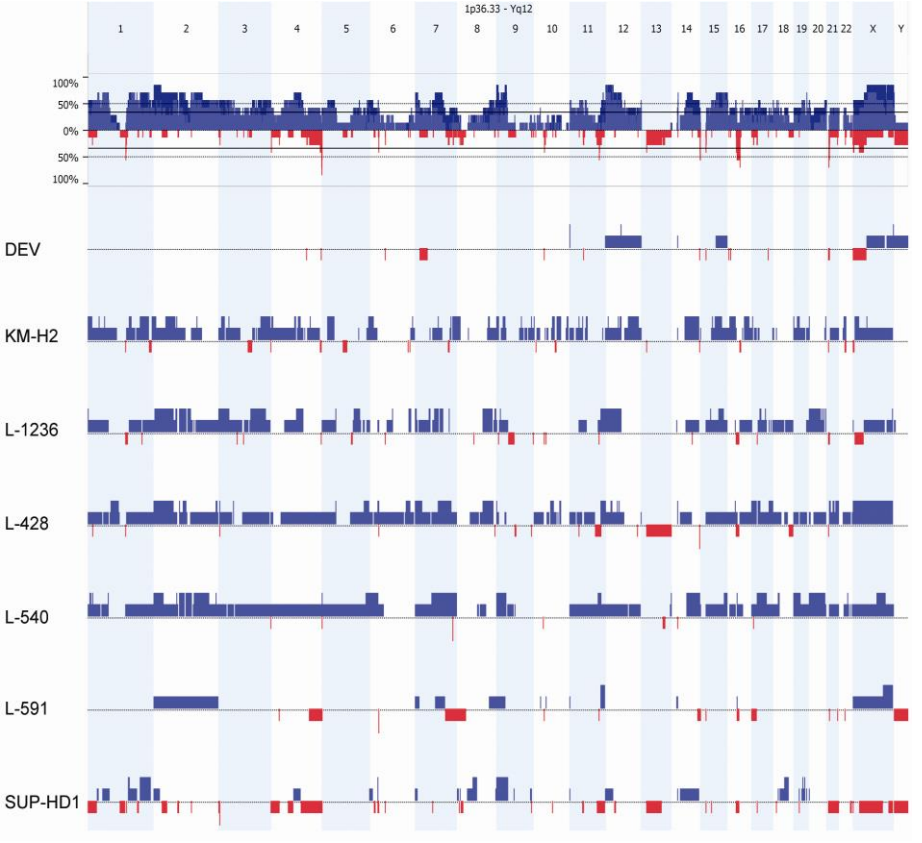
na, not applicable



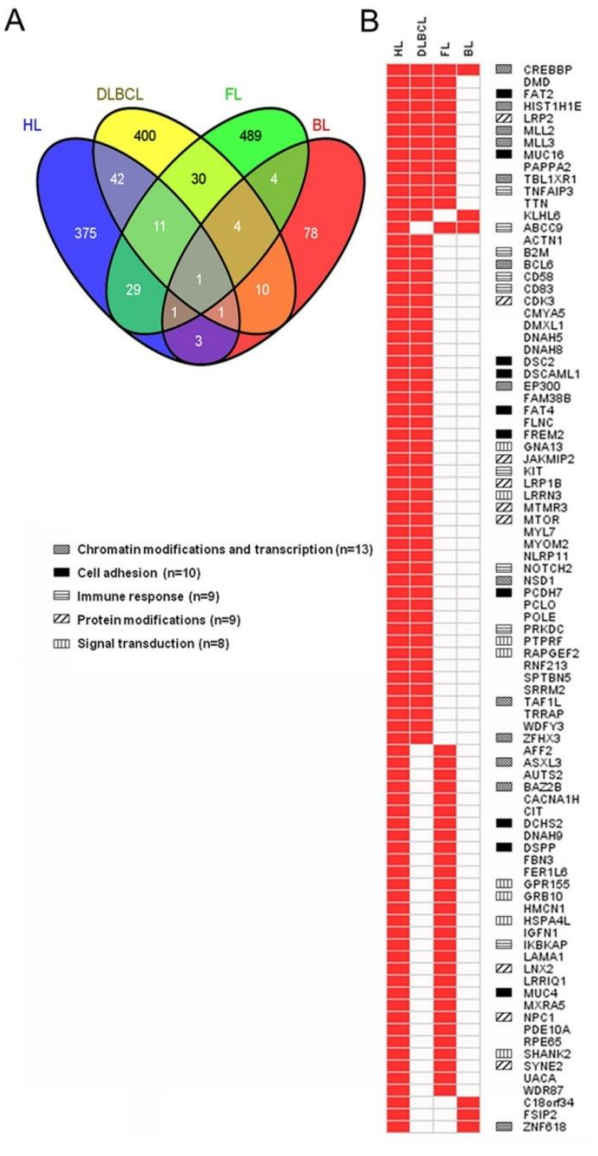

**Supplementary Figure S1.** Schematic overview of the genes mutated in at least two of the cHL cell lines. In total, 373 genes are mutated in more than 2 cHL cell lines. Of these 275 are shown in the left panel based on a high impact mutation or a SIFT score ≤0.05 in at least one cHL cell line. The 61 genes shown in the right panel include 50 genes mutated in at least 3 cHL cell lines and 11 genes carrying high impact mutations in 2 cHL cell lines. Purple boxes indicate genes with mutations that have a SNPEFF_IMPACT score "high", including stop gain, frame shift and splice site mutations. Green and grey boxes indicate mutated genes with SNPEFF_IMPACT "moderate", including missense mutations. Dark green boxes are mutations that probably are damaging based on SIFT score ≤0.05, light green boxes are mutations that are probably not damaging based on a SIFT score >0.05, grey boxes are mutations without SIFT score. The most common gene ontology groups are indicated by boxes left of the gene names.




**Supplementary Figure S2. Initial comparison of genes mutated in our WES data with genes reported to be mutated in the Broad data. (A)** Venn diagram showing that 1584 genes overlap between the Agilent capture kit and the Broad Institute. Venn diagrams to check the consistency of our WES data with the Broad data within the 1584 overlapping genes for **(B)** SUPHD1 **(C)** L540 **(D)** KMH2 **(E)** L428 and **(F)** L1236. True consistency was determined after manual inspection of the raw reads for both the WES and Broad data sets (See Supplementary Table S4).




**Supplementary Figure S3. Schematic representation of the copy number variations (CNVs) identified based on the total read counts of the WES data in all 7 HL cell lines.** The most common alterations were gains of 2p, 4q, 7p/q, 9p, 12p, 14q and 15q, and loss of 4q, 16q and 21p. In the upper panel the average frequency of the copy number gain and loss are indicated, below the copy number aberrations of each cell line is shown separately. The samples were centered by probe medians (cell lines DEV, L591, SUPHD1) or by visual selection of diploid regions (cell lines KMH2, L-1236, L-428 and L-540).




**Supplementary Figure S4. Marked overlap of genes mutated in HL as compared to DLBCL, FL and BL. (A)** Venn diagram shows the overlap between genes mutated in HL and three other NHL subtypes. **(B)** Overview of the 88 genes mutated in HL and in at least any one of the other lymphoma subtypes. Functional classification of genes is indicated by the boxes left to the gene names. 55 genes were mutated both in HL and DLBCL, 42 in HL and FL and 6 in HL and BL.

## REFERENCES


1. Kuppers R, Engert A, Hansmann ML. Hodgkin lymphoma. J Clin Invest 2012; 122(10): 3439-47.
2. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA et al. A map of human genome variation from population-scale sequencing. Nature 2010; 467(7319): 1061-73.
3. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 2010; 26(5): 589-95.
4. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010; 20(9): 1297-303.
5. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics 2009; 25(21): 2865-71.
6. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012; 6(2): 80-92.
7. Boomsma DI, Wijmenga C, Slagboom EP, Swertz MA, Karssen LC, Abdellaoui A et al. The Genome of the Netherlands: design, and project goals. Eur J Hum Genet 2013.
8. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature 2011; 476(7360): 298-303.
9. Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 2010; 26(7): 873-81.
10. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 2004; 5(4): 557-72.
11. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012; 22(3): 568-76.
12. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009; 25(16): 2078-9.
13. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4(1): 44-57.
14. Zhang J, Grubor V, Love CL, Banerjee A, Richards KL, Mieczkowski PA et al. Genetic heterogeneity of diffuse large B-cell lymphoma. Proc Natl Acad Sci U S A 2013; 110(4): 1398-403.
15. Morin RD, Mungall K, Pleasance E, Mungall AJ, Goya R, Huff RD et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. Blood 2013; 122(7): 1256-65.
16. Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet 2011; 43(9): 830-7.
17. Lohr JG, Stojanov P, Lawrence MS, Auclair D, Chapuy B, Sougnez C et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc Natl Acad Sci U S A 2012; 109(10): 3879-84.

**18.** Green MR, Gentles AJ, Nair RV, Irish JM, Kihira S, Liu CL et al. Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. Blood 2013; 121(9): 1604-11.
**19.** Richter J, Schlesner M, Hoffmann S, Kreuz M, Leich E, Burkhardt B et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. Nat Genet 2012; 44(12): 1316-20.
**20.** Gunawardana J, Chan FC, Telenius A, Woolcock B, Kridel R, Tan KL et al. Recurrent somatic mutations of PTPN1 in primary mediastinal B cell lymphoma and Hodgkin lymphoma. Nat Genet 2014; 46(4): 329-35.
**21.** Morris LG, Ramaswami D, Chan TA. The FAT epidemic: a gene family frequently mutated across multiple human cancer types. Cell Cycle 2013; 12(7): 1011-2.
**22.** Yin H, Liang Y, Yan Z, Liu B, Su Q. Mutation spectrum in human colorectal cancers and potential functional relevance. BMC Med Genet 2013; 14: 32.
**23.** Zhang J, Mullighan CG, Harvey RC, Wu G, Chen X, Edmonson M et al. Key pathways are frequently mutated in high-risk childhood acute lymphoblastic leukemia: a report from the Children's Oncology Group. Blood 2011; 118(11): 3080-7.
**24.** Diepstra A, Poppema S, Boot M, Visser L, Nolte IM, Niens M et al. HLA-G protein expression as a potential immune escape mechanism in classical Hodgkin's lymphoma. Tissue Antigens 2008; 71(3): 219-26.
**25.** Steidl C, Shah SP, Woolcock BW, Rui L, Kawahara M, Farinha P et al. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. Nature 2011; 471(7338): 377-81.

# Chapter 3

## CD58 mutations are common in Hodgkin lymphoma cell lines and loss of CD58 expression in tumor cells occurs in Hodgkin lymphoma patients who relapse

**Fazlyn Reeny Abdul Razak**, Arjan Diepstra,
Lydia Visser, Anke van den Berg

Genes & Immunity July 2016 (Doi: 10.1038/gene.2016.30)

**ABSTRACT**

CD58 is involved in immune recognition of tumor cells via binding of the CD2 receptor expressed on cytotoxic T cells. In diffuse large B cell lymphoma, mutations of the CD58 gene are reported to contribute to immune evasion of the tumor cells. We previously showed *CD58* mutations in three Hodgkin lymphoma (HL) cell lines by whole exome sequencing. In this study, we confirmed the mutations by Sanger sequencing at the DNA and RNA level and showed low levels or total loss of *CD58* mRNA expression in two of the three cell lines. CD58 protein expression as determined by flow, western blot and IHC was absent in all 3 mutated HL cell lines. In primary tissue samples loss of CD58 expression was observed in 11% of the patients who relapse. These data suggest that loss of CD58 is a potential immune escape mechanism of HL tumor cells, especially in clinically aggressive disease.

## INTRODUCTION

Hodgkin lymphoma (HL) accounts for 11% of all malignant lymphomas with around 66,000 new cases being diagnosed worldwide in 2012.[1] HL is characterized by a minority of B-cell derived tumor cells, named Hodgkin Reed-Sternberg (HRS) cells in classical (c)HL and lymphocyte predominant (LP) cells in nodular lymphocyte predominant (NLP) HL.



Our previously reported whole exome sequencing (WES) analysis of seven HL cell lines revealed recurrent mutations in multiple immune related genes.[2] The most common were *B2M* start codon mutations in L428 and DEV resulting in failure to express B2M protein. Without B2M, HLA class I cannot be properly assembled and transported to the cell membrane and this loss of HLA class I represents an immune escape mechanism for HL tumor cells. In the same study, CD58, also known as lymphocyte function-associated antigen-3, was shown to be mutated in KMH2 and DEV. CD58, a member of the immunoglobulin superfamily, is a heavily glycosylated cell-surface adhesion molecule expressed on B cells and T cells.[3,4] It is an accessory molecule that can provide a co-stimulatory signal to T cells via binding to the CD2 receptor.[5] This interaction is critical for the regulation of effector functions of T cells.[4,6] CD4+ Th2 cells form a main part of the reactive infiltrate in HL and these cells form rosettes around the HRS cells.[7] The formation of these rosettes was shown to be dependent on expression of both CD2 and CD54 on T cells and their interaction with CD58 and LFA-1 expressed on L428 cells.[8] As the HL tumor cells depend on trophic factors and growth stimuli from CD4+ T cells, CD58 expression may be necessary for tumor cell survival.[9] However, CD58 is also involved in immune recognition of tumor cells by cytotoxic T cells (CTL) and Natural Killer (NK) cells, via binding to the CD2 receptor.[10,11] Thus, loss of CD58 molecules may allow HRS cells to escape from immune recognition, especially in advanced disease when the tumor cells may have progressed to become less dependent on the reactive infiltrate and more immunogenic.[12] The aim of this study is to analyze CD58 mutations and expression in HL.

## MATERIALS AND METHODS

### *Cell lines and primary samples*

HL cell lines DEV (NLPHL, cell line developed in house, no data from DSMZ&ATCC),[13] L428 (nodular sclerosis, DSMZ No ACC 197), L1236 (mixed cellularity, DSMZ No ACC 530), KMH2 (mixed cellularity, DSMZ No ACC 8), L591 (nodular sclerosis, EBV+, DSMZ No ACC 602) and L540 (nodular sclerosis, T cell derived, DSMZ No ACC 72) were cultured in RPMI 1640 medium (Lonza Walkersville, Walkersville, MD) supplemented with 5% (L428),

20% (DEV, L540) and 10% (other cell lines) fetal calf serum, 100U/ml penicillin/streptomycin and ultraglutamine (Lonza Walkersville) in a 5% CO2 atmosphere at 37°C. SUPHD1 (lymphocyte depleted subtype, DSMZ No ACC 574) was cultured in McCoy 5A medium supplemented with 10% fetal calf serum. DLBCL cell lines SUDHL4 (DSMZ No ACC 495) was cultured in RPMI 1640 medium with 10% fetal calf serum. The origin of all cell lines was confirmed with STR DNA analysis. Mycoplasma tests consistently showed that the cell lines were not contaminated. Tissue samples of 43 HL patients without relapse and 53 HL patients with relapse (28 primary tissue samples and 31 relapse tissue samples) were retrieved from the tissue bank of the Pathology Department, University Medical Center Groningen. The study was conducted in accordance with the Declaration of Helsinki and the national Code of conduct for responsible use of human tissue (FEDERA) guidelines.

***DNA and RNA isolation and (Quantitative) RT-PCR***

DNA was isolated by salt-chloroform extraction method and RNA was isolated using TRIzol reagent. Complementary DNA (cDNA) was synthesized using Superscript II RT (Invitrogen, Carlsbad, USA) and random hexamers. qRT-PCR was performed with SYBR Green and the Lightcycler 480 (Roche, Mannheim, Germany). TBP was used as a housekeeping gene for normalization. Primers used for the amplification are listed in Table S2. Relative expression levels are expressed as $2^{-(\text{delta Cp})}$.

***Sanger sequencing***

M13 tail containing primer sets were designed to amplify the mutated region at the DNA and mRNA level (Table S2). Amplification was performed following standard protocols provided by the manufacturer. PCR products were gel purified using high pure PCR product purification kit (Roche) and send for Sanger sequencing (LGC Genomic, Teddington, UK).

***CD58 protein expression***

The expression of CD58 was analyzed by Western blot using a polyclonal anti-CD58 antibody (1:1000; AF1689, R&D Systems, Minneapolis, USA). As a loading control we analyzed expression of GAPDH (1:20000, SC47724, Santa Cruz, Heidelberg, Germany). Presence of membranous CD58 on HL cell lines was determined on the BD FACS Calibur (BD Biosciences, New Jersey, USA) using PE-labelled mouse anti-CD58 antibody (BD555921, BD Biosciences).
Immunostaining was performed with a polyclonal antibody against CD58 (AF1689, R&D Systems). After fixation in acetone for 10 minutes, cytospin slides were incubated with primary antibody (1:640) at 4°C overnight. Paraffin

embedded HL tissue sections were subjected to antigen retrieval in sodium citrate buffer (pH 6.0) in a microwave for 15 minutes. Tissue sections were incubated with primary antibody (1:80) at 4°C overnight. Positive staining was visualized after secondary antibody incubation steps with 3,3'-Diaminobenzidine (paraffin sections) or 3-amino-9-ethylcarbazole (cytospins). The tissues were examined and scored by an experienced hematopathologist without prior knowledge of patient outcomes. Positive staining was defined as clear presence of membranous staining. Variable weak background staining was scored as negative.



### *Statistical analysis*

Statistical analysis was performed by Mann Whitney U test, Fisher's exact test or Chi square test when appropriate. A p-value less than 0.05 was considered to be significant. Statistical analyses were performed with GraphPad Prism version 5.0 (GraphPad Software, San Diego, CA, USA) and SPSS statistical software (IBM SPSS Statistics 20; IBM Corporation, NY, USA).

## RESULTS AND DISCUSSION

Previously, we reported *CD58* mutations in KMH2 and DEV cells, i.e. a splice site donor and a stop gain mutation respectively.[2] A recent study by Schneider et al.[12] also reported the CD58 mutation in KMH2 and in addition showed a homozygous loss of exons 1 to 3 of the *CD58* gene in SUPHD1. Thus, *CD58* is mutated in three out of seven HL cell lines. We now confirmed the two point mutations in KMH2 and DEV by Sanger-sequencing (Figure S1). Moreover we confirmed aberrant splicing of the *CD58* transcript in KMH2 due to the splice site mutation (Figure S1D) as shown by Schneider et al.[12] and defined the size of the homozygously deleted region in SUPHD1 (Figure S1). As DEV is a cell line derived from nodular lymphocyte predominant HL, this finding showed that *CD58* alterations may also be important in this HL subtype.

In comparison to the HL cell lines with wild type *CD58*, the mRNA levels were low or absent in SUPHD1 and KMH2 cells (Figure 1B). In DEV cells the levels were comparable to the other HL cell lines. CD58 membrane expression as determined by flow (Figure 1C; Figure S2A) showed lack or low expression in all three mutated HL cell lines. These findings were confirmed by Western blot (Figure 1D). Immunostaining of the HL cell lines revealed no staining in the mutated cell lines, while the other cell lines showed strong staining (Figure S2B). These data confirm presence of *CD58* gene mutations with concurrent loss of expression of the CD58 protein in 3 out of 7 HL cell lines.

Schneider et al.[12] identified heterozygous deletions of *CD58* by FICTION analysis in 3 out of 13 primary cHL cases. Sequencing analysis of the *CD58* gene on DNA isolated from microdissected HRS cells of 10 cHL cases did not reveal any mutations. Reichel et al.[14] showed *CD58* gene deletions in 2 out of 10 flow-isolated HRS cells cHL cases. Based on these studies and on our data, it is obvious that alterations of the *CD58* gene are present in primary HRS cells of patients that do experience a relapse, albeit at a low percentage. Unfortunately, we were not able to determine presence of mutations in our cases with loss of CD58 expression, as we do not have frozen tissue samples or cell suspensions available (see below).

In addition to mutations in *CD58*, various other immune associated genes have been shown to be mutated in HL. The CSF2RB gene, encoding the common β chain (CD131) shared by the interleukin-3 (IL-3), granulocytic macrophage colony-stimulating factor (GM-CSF) and IL-5 receptors, was mutated in 4 out of 7 HL cell lines.[2] STXBP2 required for exocytosis of perforin-containing lytic granules was mutated in SUPHD1, L1236 and DEV. Moreover, a bi-allelic mutation was reported in STXBP2 in an Epstein-Barr virus positive cHL case by Machaczka et al.[15] We reported *B2M* gene mutations affecting the ATG start codon in L428 and DEV cells.[2] A recent study by Reichel et al.[14] reported that *B2M* is the most frequently mutated gene in cHL. Mutations were strongly associated with the nodular sclerosis subtype, younger age and better overall survival. B2M is essential for proper transport of HLA class I to the cell membrane and loss of B2M results in loss of membranous HLA-class I expression.

Diffuse large B-cell lymphomas (DLBCL) are also characterized by mutations of *CD58* and *B2M*. Loss of CD58 expression is associated with loss of HLA class I in DLBCL[11] and this co-loss is crucial to avoid triggering of NK cell activation by CD58. We showed loss of CD58 protein expression in SUPHD1, DEV and KMH2 cells as a consequence of mutations in the *CD58* gene in this study. In a previous study we showed *B2M* mutations and loss of HLA class I expression in L428 and DEV cells.[12] This indicates that L428 cells lost both CD58 and B2M, making them vulnerable to NK cell killing, but these cells apparently survived, possibly due to the protection from NK cells by the rosetting CD4+ Th2 cells.[8] So, in contrast to DLBCL, loss of CD58 expression is not restricted to HLA class I negative HL cell lines.

We next studied CD58 expression in primary tissue samples of 43 HL patients without relapse (average follow-up time 6.4 years (0.04 – 16.6 years) (Table 1). Consistent with earlier studies,[9,16] we showed a strong CD58 staining in HRS cells of all HL patients. CD58 expression may be important for HRS cells to maintain their capacity as antigen presenting cells and to interact with CD4+ T cells that offer growth and survival to the tumor cells.[9,16] As HL cell lines are derived from refractory or relapsed HL disease with extensive extranodal involvement, it might be that loss of CD58 expression is a late event in HL progression and may occur when HRS cells want to become independent of the T-cell infiltrate for survival. At that point loss of HLA class I and CD58 might be a prerequisite for survival of tumor cells as the tumor cells have accumulated mutations that can potentially lead to presentation of highly immunogenic peptides that can activate cytotoxic T cell responses. We next studied CD58 expression in 53 relapsed HL patients. For 22 patients we only had primary diagnostic material, for 25 patients we only had relapse tissue and for 6 patients we had both (Table S1). In contrast to the consistent strong staining pattern of CD58 on HRS cells of patients who did not develop a relapse, loss of CD58 staining was observed in 4 out of 28 primary and in 3 out of 31 relapse tissue samples. These 7 negative cases were derived from 6 patients who did relapse. One of the 6 paired cases showed a positive staining of HRS cells in the primary tissue sample and no staining of the HRS cells in the relapse tissue samples. Tissue samples of all other relapsed patients showed a strong staining. The percentage of cases with CD58 loss in relapsed HL patients was significantly higher than in the HL patients without relapse. These staining data showed that loss of CD58 on tumor cells is restricted to HL patients who do experience relapse (Table 1, Figure S3). Four of the six CD58 negative cases, also stained negative for HLA class I, while the other 2 cases were not scorable, because of an insufficient amount of tissue. Due to the low number of CD58 negative cases, we cannot conclude whether loss of CD58 is specific for HLA class I negative cases, similar to the findings in DLBCL.[11]

In conclusion, our results indicate that mutations in *CD58* and loss of CD58 expression are common in HL derived cell lines and that loss of CD58 expression in tumor cells is associated with relapse of disease.

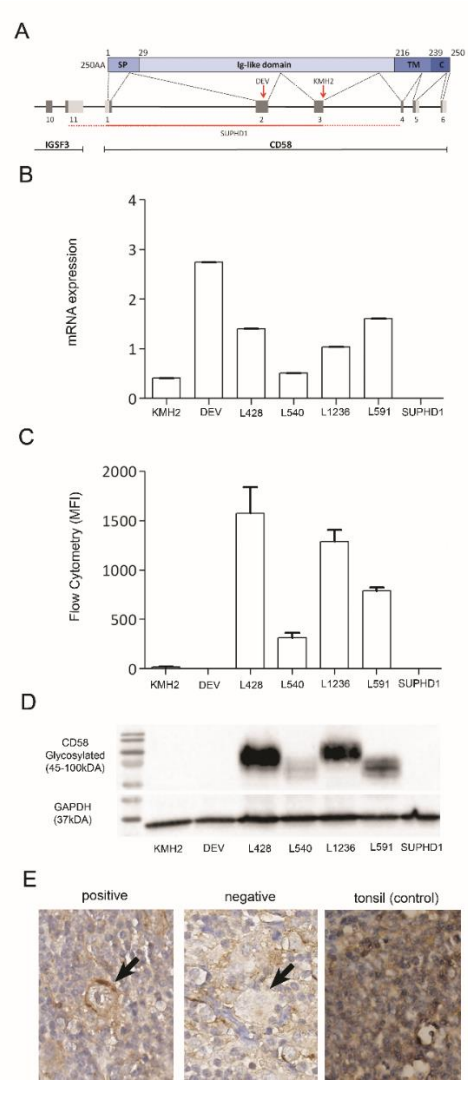



**Figure 1. CD58 expression in HL. (A)** Schematic diagram of the *CD58* gene and protein. Red arrows indicate the position of the stop gain mutation in DEV and the splice donor site mutation in KMH2. The red line indicates the homozygously deleted region in SUPHD1. The solid line represents the minimal deleted region and the dotted line represents the region for which loss is uncertain. The second allele most likely carriers a larger deletion. DEV and KMH2 have homozygous mutations. **(B)** The bar graph below shows the qRT-PCR quantification of the relative CD58 mRNA expression (mean ± SD); **(C)** Flow cytometric analysis (mean fluorescence intensity, MFI ± SD) of membraneous CD58 in 7 HL cell lines using PE-labelled mouse anti-CD58 antibody (BD555921, BD Biosciences). **(D)** Protein expression as determined by Western blot using a polyclonal anti-CD58 antibody (1:1000; AF1689, R&D Systems, Minneapolis, USA) and GAPDH (1:20000, SC47724, Santa Cruz, Heidelberg, Germany). The molecular weight of the glycosylated CD58 protein is 45-100kDA depending on the extent of glycosylation on the extracellular domain. **(E)** Immunostaining of paraffin embedded HL tissue sections using a polyclonal antibody against CD58 (1:80, AF1689, R&D Systems). 6 out of 53 HL cases with relapse have diminished staining compared to HL cases without relapse. Shown are representative examples of positive and negative staining. Some cells in the microenvironment are positive as CD58 is expressed on B cells and T cells. Tonsil tissue was used as a positive control; the germinal center is shown. Magnification at 80X.

**Table 1.** Distribution of age, follow up time, gender, histological subtype and CD58 staining in HL patients without relapse and HL patients with relapse.

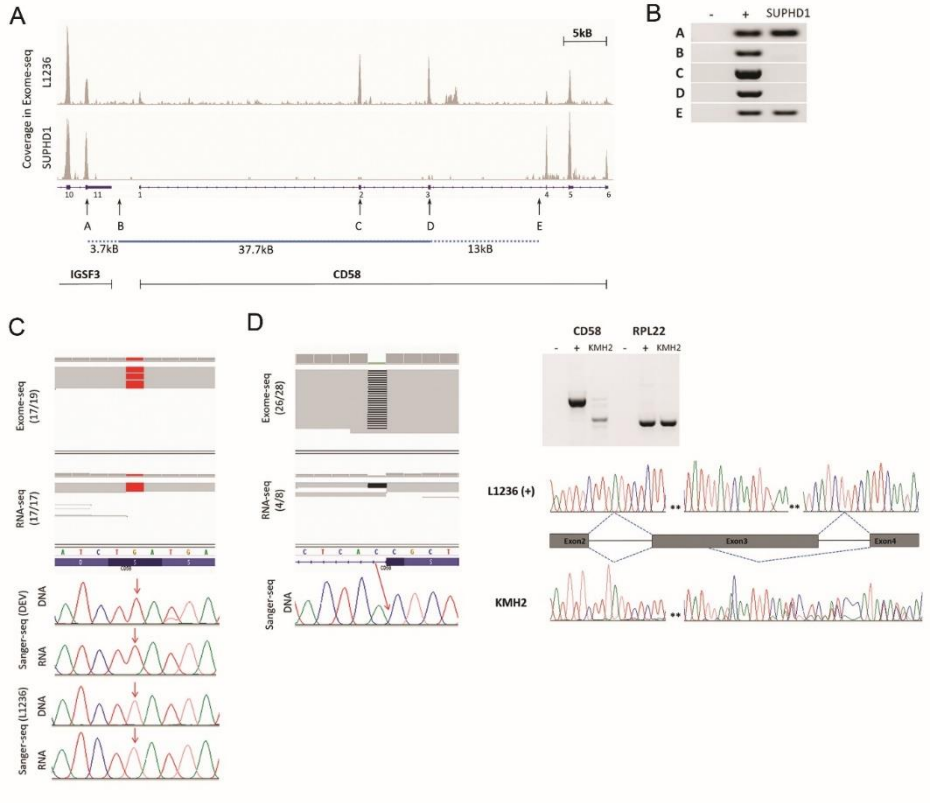
| Characteristic | HL patients without relapse (n=43) | HL patients with relapse (n=53) | p-value |
|---|---|---|---|
| Median age (range) | 26 (9-75) | 30 (5-62) | ns‡ |
| Median follow up time (years) | 6.4 (0.04-16.6) | 1.5 (0.4-20.6) | 0.0001‡ |
| Gender | | | |
| Male | 23 (53%) | 29 (55%) | ns† |
| Female | 20 (47%) | 24 (45%) | |
| Histological subtype | | | |
| cHL | | | |
| Nodular sclerosis | 29 (67%) | 42 (79%) | ns* |
| Mix cellularity | 5 (12%) | 8 (15%) | |
| Lymphocyte rich | 2 (5%) | 1 (2%) | |
| NOS | 3 (7%) | - | |
| NLPHL | 4 (9%) | 2 (4%) | |
| CD58 staining | | | |
| Positive | 43 (100%) | 47 (89%) | 0.03† |
| Negative | - | 6 (11%) | |

‡Mann Whitney U test, †Fisher's exact test, *Chi square test. HL: Hodgkin lymphoma; cHL: classical Hodgkin lymphoma; NOS: cHL not otherwise specified; NLPHL: nodular lymphocyte predominant Hodgkin lymphoma

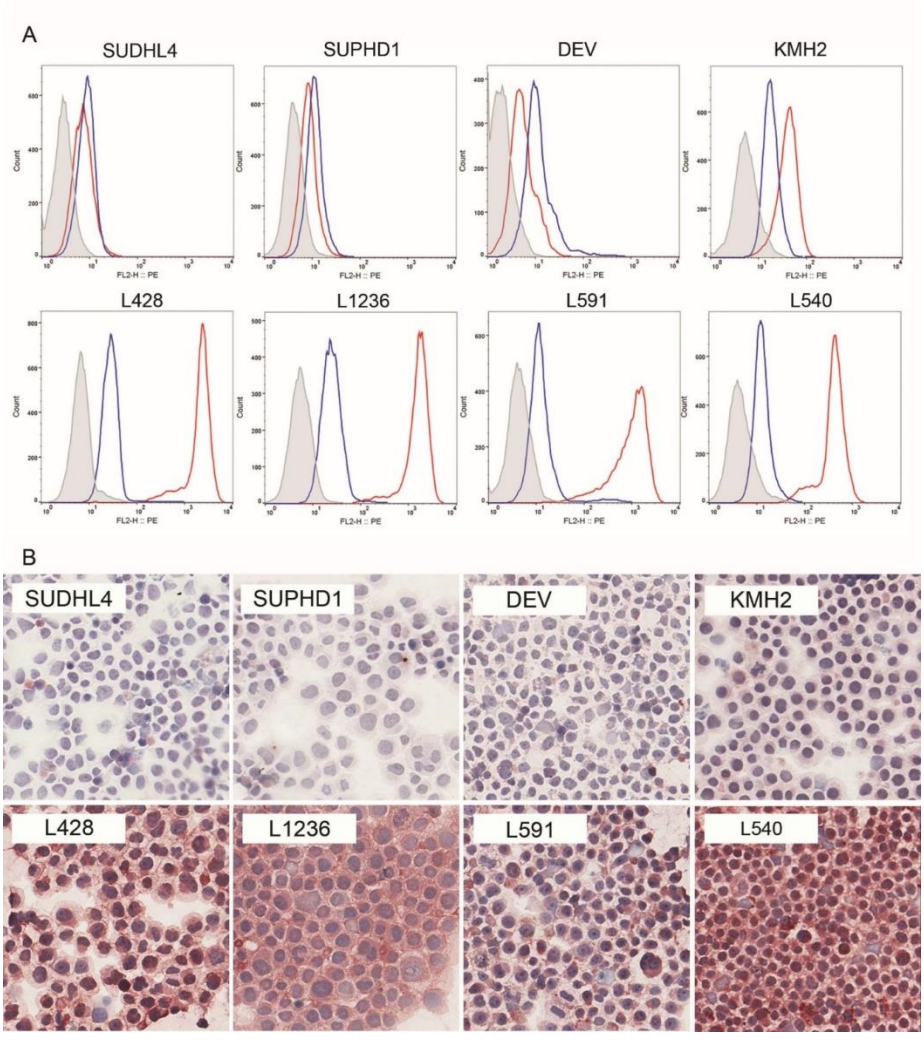
**ACKNOWLEGMENTS**

This work was supported by grants from the Abel Tasman Talent Program of the University Medical Center Groningen.

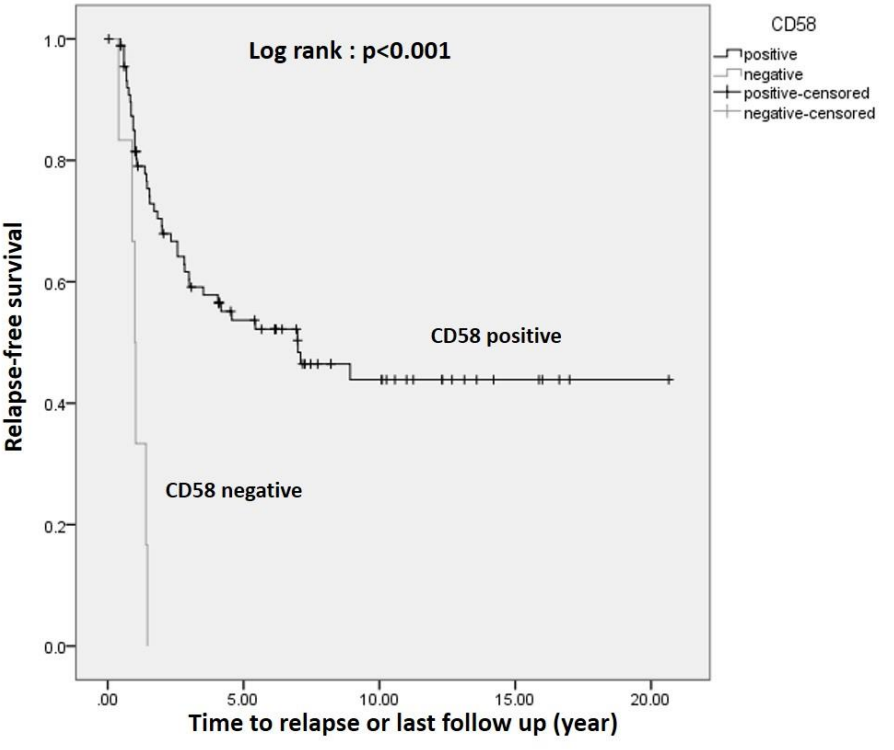
## Supplementary Information




**Supplementary Figure S1. CD58 mutations in HL cell lines.** (**A**) Homozygous deletion of CD58 in SUPHD1. We inspect the coverage by using the IGV and run PCR with multiple primer set to validate the deletion. The arrows indicate the primer position. The solid line represent homozygously deleted region and dotted line represent the borders of the deletion for which loss is uncertain. (**B**) PCR validation for deletion in SUPHD1 revealed that the size of this single deletion is between 37.7kB (the distance between primer sets B and D) and 54.4kB (the distance between primer set A and E). (**C**) Validation of the stop gain mutation in DEV at the DNA and RNA level by PCR and Sanger-sequencing. (**D**) Validation of the splice donor site mutation in KMH2 at the DNA level and analysis of the effect on splicing by RT-PCR and Sanger sequencing.




**Supplementary Figure S2. CD58 expression in HL cell lines. (A)** CD58 protein expression in HL cell lines as detected by FACS analysis. Gray-filled: blank; Blue line: isotype control; Red line: samples incubated with CD58 antibody. (**B**) CD58 immunohistochemistry of HL cell lines. SUDHL4 cells were used as negative control. CD58 was absent in SUPHD1 with a homozygous deletion, DEV with stop gain mutation and KMH2 with splice donor site mutation. CD58 expression was detected in the other HL cell lines, i.e. L428, L1236, L591 and L540.




**Supplementary Figure S3. Kaplan–Meier survival curve for relapse-free survival of Hodgkin lymphoma patients (n=96) who were CD58 positive (n=90) or CD58 negative (n=6).** CD58 negative patients had a significantly poorer relapse-free survival than CD58 positive patients (log rank test p<0.001).

**Supplementary Table S1.** Distribution of age, gender, histological subtype and CD58 staining in HL patients without relapse and HL patients with relapse.

| Characteristic | Primary tissue of HL patients without relapse (n=43) | Tissue of HL patients with relapse (n=53*) | |
|---|---|---|---|
| | | Primary tissue (n=28) | Relapse tissue (n=31) |
| Median age (range) | 26 (9-75) | 31.5 (19-62) | 27 (5-59) |
| Gender | | | |
| Male | 23 (53%) | 15 (54%) | 18 (58%) |
| Female | 20 (47%) | 13 (46%) | 13 (42%) |
| Histological subtype | | | |
| cHL | | | |
| Nodular sclerosis | 29 (67%) | 21 (75%) | 25 (81%) |
| Mix cellularity | 5 (12%) | 6 (21%) | 4 (13%) |
| Lymphocyte rich | 2 (5%) | - | 1 (3%) |
| NOS | 3 (7%) | - | - |
| NLPHL | 4 (9%) | 1 (4%) | 1 (3%) |
| CD58 staining | | | |
| Positive | 43 (100%) | 24 (86%) | 28 (90%) |
| Negative | - | 4 (14%) | 3 (10%) |

*6 patients with both primary and relapse tissue. 4 patients have positive staining of HRS cells in both primary and relapse tissues samples, 1 patient showed no staining of HRS cells in both primary and relapse tissue samples and 1 patient showed a positive staining of HRS cells in the primary tissue sample and no staining of the HRS cells in the relapse tissue sample.

**Supplementary Table S2.** Primer sequences for PCR and qPCR.

| Primer | Forward (5' → 3') | Reverse (5' → 3') | Expected product size (bp) |
|---|---|---|---|
| PCR | | | |
| SUPHD1_A (DNA) | CCGGAACTCCAGCAAGAACT | CCTGGATGGATACTGAGAAC | 120 |
| SUPHD1_B (DNA) | TGCTGGCTTGGCCAGGAGTT | TGCACGAAGGACATAGAAGG | 185 |
| SUPHD1_C (DNA) | GTTGGTGCTCCATGAATGTC | CTAGCTCCTCTATACCAGAA | 656 |
| SUPHD1_D (DNA) | GAGGTTGCGAGCTCTTCATA | ACCACATCTGTGGTCTGAAA | 400 |
| SUPHD1_E (DNA) | AATGGAGGCACTGAGAGGTT | TCTCATACTGCCAGTGGAAG | 186 |
| DEV (DNA & cDNA) | TTCCATGTACCAAGCAATGT | GAACTTCATGGTATCAGTAA | 222 |
| KMH2 (DNA) | CAACAGCCATCGAGGACTTA | AGAGCGAGACTCTGTCTCAA | 372 |
| KMH2 (cDNA) | GTGTCAGGTAGCCTCACTAT | AGTGCATATCTGTGTCTTGA | 398 |
| RPL22 (DNA & cDNA) | TCGCTCACCTCCCTTTCTAA | TCACGGTGATCTTGCTCTTG | 250 |
| qPCR | | | |
| CD58 | AAGACACAGATATGCACTTA | TCTGTCACATTTCAGAATAC | 91 |
| TBP | GCCCGAAACGCCGAATAT | CCGTGGTTCGTGGCTCTCT | 73 |

## REFERENCES


1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013; 49(6): 1374-403.
2. Liu Y, Abdul Razak FR, Terpstra M, Chan FC, Saber A, Nijland M et al. The mutational landscape of Hodgkin lymphoma cell lines determined by whole-exome sequencing. Leukemia 2014; 28(11): 2248-51.
3. Dustin ML, Selvaraj P, Mattaliano RJ, Springer TA. Anchoring mechanisms for LFA-3 cell adhesion glycoprotein at membrane surface. Nature 1987; 329(6142): 846-8.
4. Springer TA, Dustin ML, Kishimoto TK, Marlin SD. The lymphocyte function-associated LFA-1, CD2, and LFA-3 molecules: cell adhesion receptors of the immune system. Annu Rev Immunol 1987; 5: 223-52.
5. Dengler TJ, Hoffmann JC, Knolle P, Albert-Wolf M, Roux M, Wallich R et al. Structural and functional epitopes of the human adhesion receptor CD58 (LFA-3). Eur J Immunol 1992; 22(11): 2809-17.
6. Springer TA. Adhesion receptors of the immune system. Nature 1990; 346(6283): 425-34.
7. Poppema S, Potters M, Visser L, van den Berg AM. Immune escape mechanisms in Hodgkin's disease. Ann Oncol 1998; 9 Suppl 5: S21-4.
8. Sanders ME, Makgoba MW, Sussman EH, Luce GE, Cossman J, Shaw S. Molecular pathways of adhesion in spontaneous rosetting of T-lymphocytes to the Hodgkin's cell line L428. Cancer Res 1988; 48(1): 37-40.
9. Ellis PA, Hart DN, Colls BM, Nimmo JC, MacDonald JE, Angus HB. Hodgkin's cells express a novel pattern of adhesion molecules. Clin Exp Immunol 1992; 90(1): 117-23.
10. Warren HS, Smyth MJ. NK cells and apoptosis. Immunol Cell Biol 1999; 77(1): 64-75.
11. Challa-Malladi M, Lieu YK, Califano O, Holmes AB, Bhagat G, Murty VV et al. Combined genetic inactivation of beta2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. Cancer Cell 2011; 20(6): 728-40.
12. Schneider M, Schneider S, Zuhlke-Jenisch R, Klapper W, Sundstrom C, Hartmann S et al. Alterations of the CD58 gene in classical Hodgkin lymphoma. Genes Chromosomes Cancer 2015; 54(10): 638-45.
13. Poppema S, Visser L, de Jong B, Brinker M, Atmosoerodjo J, Timens W. The typical Reed-Sternberg phenotype and Ig gene rearrangement of Hodgkin's disease derived cell line ZO indicating a B-cell origin. Recent Results Cancer Res 1989; 117: 67-74.
14. Reichel J, Chadburn A, Rubinstein PG, Giulino-Roth L, Tam W, Liu Y et al. Flow-sorting and exome sequencing reveals the oncogenome of primary Hodgkin and Reed-Sternberg cells. Blood 2014.
15. Machaczka M, Klimkowska M, Chiang SC, Meeths M, Muller ML, Gustafsson B et al. Development of classical Hodgkin's lymphoma in an adult with biallelic STXBP2 mutations. Haematologica 2013; 98(5): 760-4.
16. Jacob MC, Agrawal S, Chaperot L, Giroux C, Gressin R, Le Marc'Hadour F et al. Quantification of cellular adhesion molecules on malignant B cells from non-Hodgkin's lymphoma. Leukemia 1999; 13(9): 1428-33.

# Chapter 4

## *CSF2RB* gene mutations in Hodgkin lymphoma

**Fazlyn Reeny Abdul Razak**, Arjan Diepstra, Emmie de Vries, Lydia Visser, Anke van den Berg

Work in progress

**ABSTRACT**

The percentage of Hodgkin and Reed-Sternberg (HRS) tumor cells in Hodgkin lymphoma (HL) is very low and they reside within an abundant inflammatory microenvironment. Chemokines and cytokines produced by HRS cells and by infiltrating cells shape the environment and provide proliferative and survival signals to the HRS cells. The aim of this study was to confirm the previously identified mutations of *CSF2RB* in four out of seven HL cell lines and to identify the functional relevance of these mutations. *CSF2RB* encodes for the common β chain shared by the interleukin-3 (IL-3), interleukin-5 (IL-5) and granulocytic macrophage colony-stimulating factor (GM-CSF) receptors. CSF2RB signals through the JAK2/STAT5 pathways to promote survival, proliferation and differentiation. We validated the cell line mutations by Sanger sequencing and RNA sequencing. Two missense mutations (V212I and A715S) in KMH2, a missense (V524M) and frame shift mutation (c.2500_2501del) in SUPHD1, a frame shift mutation (c.2198del) in L1236 and a stop gain mutation (Q809*) in DEV cells were confirmed at the DNA and mRNA levels. We were unable to determine whether the two mutations present in KMH2 and SUPHD1 were present on the same parental allele or not. By flow cytometry, CSF2RB levels were slightly higher in the cell lines with a mutation as compared to the HL cell lines with wild type CSF2RB. SUPHD1 and L428 showed high IL3RA expression levels compared to the other cells at the mRNA and protein level. *IL5RA* mRNA expression was detected in SUPHD1, L1236 and L428 and not in the other 4 HL lines. CSF2RA expression levels were very low or undetectable in all seven HL cell lines. For cytokines, IL-3 levels were undetectable in culture medium of HL cell lines, IL-5 was detectable only in L428 and GM-CSF was detected in 4 HL cell lines. No effects were observed upon stimulation with any of the three cytokines. In conclusion, we confirmed the mutations previously reported by whole exome sequencing. However, we did not see an effect on growth upon exogenous cytokine addition.

## INTRODUCTION

Hodgkin Lymphoma (HL) is a B cell derived malignancy characterized by a minority of tumor cells, known as Hodgkin Reed-Sternberg (HRS) cells. The abundant reactive background is composed of a wide variety of inflammatory cells with a predominance of T cells. Chemokines and cytokines produced by HRS cells and by the infiltrating cells shape the environment and provide proliferative and survival signals to the HRS cells. Despite the loss of B-cell receptor (BCR), the master regulator of B cell fate, HRS cells survive due to constitutive activation of a number of pathways that regulate tumor cell survival.[1,2]



Whole exome sequencing (WES) of seven HL cell lines revealed recurrent mutations in *CSF2RB*.[3] *CSF2RB* encodes for the common β chain shared by the interleukin-3 (IL-3), interleukin-5 (IL-5) and granulocytic macrophage colony-stimulating factor (GM-CSF). CSF2RB is co-expressed on the cell surface together with the unique α chains of either IL-3, IL-5 or GM-CSF, i.e. IL3RA, IL5RA and CSF2RA. This heterodimeric receptor binds IL-3, IL-5 or GM-CSF with high affinity and induces downstream signaling by phosphorylation of tyrosine residues in CSF2RB. This induces activation of Janus kinase 2 (JAK2) and phosphorylation of STAT5. The cytokines that share the common β chain have different roles in HL. HRS cells express IL3RA, but not IL-3 and addition of IL-3 to HL cell line culture medium promotes cell growth.[4] Multiple cells present in the microenvironment of cHL produce IL-3 such as activated T lymphocytes, mast cells and eosinophils. Thus, IL-3 probably acts as a paracrine growth factor for the HRS cells.[4,5] IL-5 is essential for eosinophil growth and differentiation[6-8] and has been identified in primary HRS cells especially in cases with high numbers of eosinophils[8]. IL-5 is probably not a growth factor for cHL, since IL-5 antibodies do not block cell growth in HL cell lines.[7] GM-CSF is produced by HL cell lines and is an essential factor for the differentiation of granulocytes, macrophages and eosinophils.[6,9,10]

In this study we aimed to confirm the previously identified mutations of *CSF2RB* in four HL cell lines. In addition, we present preliminary data on the functional relevance of these mutations by analyzing expression of CSF2RB and the three α chains and by determining the effect of cytokine stimulation on the growth of HL cells.

## MATERIALS AND METHODS

### *Cell lines and primary samples*

DEV (NLPHL, cell line developed in house),[11] L428 (nodular sclerosis, DSMZ No ACC 197), L1236 (mixed cellularity, DSMZ No ACC 530), KMH2 (mixed cellularity, DSMZ No ACC 8), L591 (nodular sclerosis, EBV+, DSMZ No ACC 602) and L540 (nodular sclerosis, T cell derived, DSMZ No ACC 72) were cultured in RPMI 1640 medium (Lonza Walkersville, Walkersville, MD) supplemented with 5% (L428), 20% (DEV, L540) or 10% (other cell lines) fetal calf serum, 100U/ml penicillin/streptomycin and ultraglutamine (Lonza Walkersville) in a 5% CO2 atmosphere at 37°C. SUPHD1 (lymphocyte depleted subtype, DSMZ No ACC 574) was cultured in McCoy 5A medium supplemented with 10% fetal calf serum. The origin of all cell lines was confirmed with STR DNA analysis 2 to 3 times per year. Mycoplasma tests were performed routinely to exclude contamination of the cell lines.

### *DNA and RNA extraction and purification*

DNA was isolated using the salt-chloroform extraction method and RNA was isolated using TRIzol (Life technologies, Carlsbad, USA) following standard protocol provided by the manufacturer. DNA and RNA concentrations and purity were measured on a Nanodrop ND-1000 Spectrophotometer (Nanodrop Technologies, Wilmington, USA). Only good quality DNA and RNA samples were used for further analysis.

### *Sanger sequencing and RNA sequencing*

Primer sets were designed to amplify the region containing the known *CSF2RB* mutations at the DNA and mRNA level (Table 1). Amplification reactions were performed using Taq DNA polymerase kit (Invitrogen, Carlsbad, USA) and Qiagen long-range PCR kit (Qiagen, Venlo, The Netherlands) following standard protocols provided by the manufacturer. PCR products were gel purified using high pure PCR product purification kit (Roche, Mannheim, Germany) and sent for Sanger sequencing (LGC Genomic, Teddington, UK). RNA-seq data from previous publications was used (Steidl et al.[12] and Twa et al.[13], personal communication).

### *cDNA synthesis and (Qualitative) RT-PCR*

Complementary DNA (cDNA) was synthesized from 500ng of total RNA using Superscript II RT (Invitrogen) and random hexamers in a final volume of 20µl. Quantitative (q)RT-PCR was performed in triplicate in a volume of 10ul using SYBR Green master mix and the Lightcycler 480 (Roche, Mannheim,

Germany). TBP was used as a housekeeping gene for normalization. Relative expression levels are expressed as $2^{-(\text{delta Cp})}$. Primers used for the amplification are listed in Table 1.

***CSF2RB, CSF2RA, IL5RA and IL3RA protein expression***

The presence of membranous CSF2RB, IL3RA, IL5RA and CSF2RA on HL cell lines was determined on the BD FACS Calibur (BD Biosciences, New Jersey, USA) using PE-labeled mouse anti-CSF2RB antibody (ab93509, Abcam), PE mouse IgG1 isotype control (IQP-191R, IQP Products QV, Groningen, The Netherlands), APC mouse anti-IL3RA (BD560087, BD Biosciences), APC mouse IgG2a κ isotype control (BD555576, BD Biosciences), rabbit anti-IL5RA (PA5-25159, Thermo Fisher), rabbit IgG monoclonal isotype control (ab172730, Abcam), Alexa Fluor® 647 mouse anti-CSF2RA (BD564046, BD Biosciences) and Alexa Fluor® 647 mouse IgG1 κ isotype control (BD557714, BD Biosciences).



***IL-3, IL-5 and GM-CSF levels in culture supernatant of HL cell lines***

$2x10^5$ cells per ml were cultured in triplicate in RPMI 1640 with FCS and supernatants were collected after 24 hours. IL-3, IL-5 and GM-CSF levels were measured by ELISA kits, IL-3 (DY203-05, R&D Systems, Minneapolis, MN, USA), IL-5 (IL-3 (DY205-05, R&D Systems) and GM-CSF (IL-3 (DY215-05, R&D Systems) according to the instructions provided by the manufacturer. RPMI 1640 supplemented with 10% FCS was used as the baseline for the levels of secreted cytokines.

***Alamar blue test***

Cellular growth after stimulation and inhibition was evaluated by alamar blue tests (AbD Serotec, Oxford, UK). Cells were seeded at 10,000 - 20,000 cells/ml in 96-well plates in serum free medium and 0, 2, 5 and 10ng/mL of each cytokine, human IL-3 (203-IL, R&D Systems) and human IL-5, (205-IL, R&D Systems). After 44 hours, 10 µL of alamar blue was added to each well. The fluorescence was measured at 48 hours, 52 hours and 68 hours with an excitation of 560 nm and emission at 590 nm.

**RESULTS AND DISCUSSION**

Previously, we reported heterozygous *CSF2RB* mutations in KMH2, SUPHD1, DEV and L1236 cells by WES[3] (Figure 1A). We found two missense mutations (V212I and A715S) in KMH2, a missense (V524M) and a frame shift mutation (c.2500_2501del) in SUPHD1, a frame shift mutation (c.2198del) in L1236 and

a stop gain mutation (Q809*) in DEV cells. We confirmed these mutations at the DNA and mRNA levels by Sanger sequencing (Figure 1B). Since KMH2 and SUPHD1 each contained 2 mutations we tried to determine whether these were present at the same parental allele or not. Despite several attempts using normal and long-range PCR approaches we were unable to amplify the region covering both mutations in KMH2 and SUPHD1 at the DNA or RNA level, so this remains undetermined.

Analysis of published data on flow-sorted HRS cells of 10 cHL cases revealed *CSF2RB* mutations in 3 cases, i.e. 2 frameshift mutations (c.2500_2501del and c.2360dup) and 1 stop gain mutation (Q817*)[14] (Figure 1A). One of the frameshift mutations was identical to the frame shift mutation observed in SUPHD1. This deletion in exon 14 (c.2500_2501del) overlaps with the region that has been shown to negatively regulate growth. These findings indicate that *CSF2RB* mutations are present not only in HL cell lines but also in HRS cells of primary cHL cases.

In 1996, Freeburn et al.[15] showed that *CSF2RB* is commonly mutated in acute myeloid leukemia (AML). However, they suggested that these somatic mutations were unlikely to contribute to the pathogenesis of AML because all variants were also identified as known polymorphisms. The 8 mutations identified in HL were not present in dbSNP (version 148). The Catalogue of Somatic Mutations in Cancer (COSMIC) lists 15 missense mutations and 3 nonsense mutations in the *CSF2RB* gene (Figure 1A). *CSF2RB* mutations have been reported at frequencies below 1%, with the highest frequencies in skin cancer (0.5%), cervical carcinoma (0.3%), intestinal adenocarcinoma (0.3%) and bladder cancer (0.3%) (Figure 1A). Recently, Watanabe et al.[16] found seven *CSF2RB* mutations in 449 primary leukemic samples. In addition, they reported one mutation previously detected as a germ line variant,[17] i.e. R461C (in T-ALL) (Figure 1A). Functional testing of these mutations revealed growth factor-independent growth of T-AML only of the R461C variant. This variant caused accumulation and phosphorylation of CSF2RB, as well as constitutive activation of the JAK/STAT pathway.[16] By using a PCR-based random mutagenesis approach combined with a retroviral expression system, the oncogenic potential of different CSF2RB mutations was tested (I374N, Q375P, Y376N, W383R, L399P, L445Q, V449E, A459D, R461C and H544R).[18,19] All mutants were able to convert a growth factor-dependent hematopoietic cell line into a growth factor independent cell line. So far, the *CSF2RB* mutation frequency appears to be the highest in HL, with 4 out of 7 HL cell lines and 3 out of 10 primary cases, reaching over 30%.

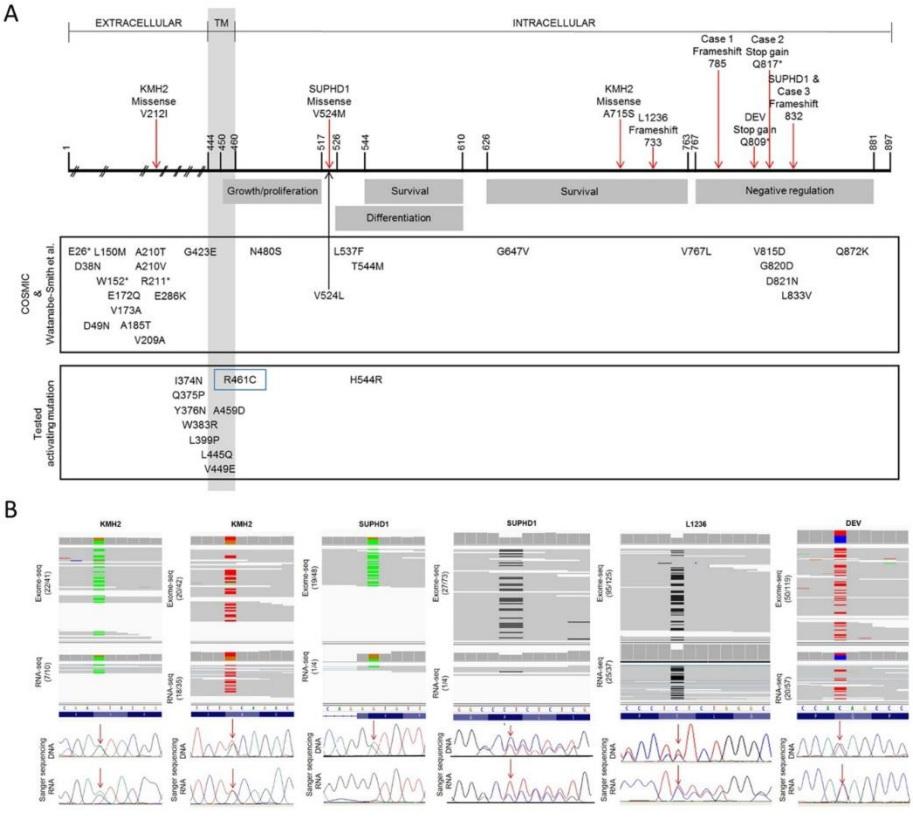



**Figure 1. CSF2RB mutations in HL cell lines. (A)** Schematic diagram of CSF2RB. Red arrows indicate the position of mutations in 4 HL cell lines (KMH2, SUPHD1, L1236 and DEV) with 2 mutations in KMH2 and SUPHD1 cell lines and also of mutations found in 3 primary HL cases by Reichel et al.[14]. The first box indicates the position of possibly pathogenic mutations reported in COSMIC and a recent study by Watanabe et al.[16]. The second box indicates the position of mutation with proven activating effects. The mutation in the blue box was reported to be a germline mutation. **(B)** Results of the previously reported WES study and the validation by RNA-seq (upper panel) and validation of the mutations by Sanger-seq at the DNA and RNA level.

CSF2RB mRNA expression levels were slightly higher in the cell lines with a mutation as compared to the HL cell lines with wild type CSF2RB (Figure 2A). Analysis of membranous CSF2RB protein expression as determined by flow revealed similar levels in all HL cell lines (Figure 2B). On the cell surface, CSF2RB forms a cytokine receptor complex with one or more α chains (IL3RA, IL5RA and CSF2RA). Depending on the α chains, these complexes have a high affinity for IL-3, IL-5 or GM-CSF. To further study the relevance of

CSF2RB, we studied expression levels of the 3 cytokines and the 3 α chains of these cytokine receptors. L428 and SUPHD1 showed high IL3RA expression levels compared to the other cell lines at both the mRNA (Figure 3A) and protein level (Figure 3B). L1236 showed lower protein levels despite high mRNA levels. Aldinucci et al.[4] also reported high expression of IL3RA in L428. In addition, they reported IL3RA expression on HRS cells of 19 HL cases by immunostaining of frozen tissue sections and cytospots of HL cell suspensions. Bosshart et al.[20] also reported surface expression of IL3RA on primary HRS cells of lymph nodes involved with NS and lymphocyte rich subtypes of cHL. IL-3 cytokine levels were undetectable in culture medium of HL cell lines (Figure 3C). This is in agreement with previous studies in HL cell lines[4,6] and in 16 primary HL cases[9]. Aldinucci et al.[4] reported enhanced growth of HL cell lines upon exogenous IL-3 using a clonogenic growth assay and 3H-thymidine incorporation assay. In contrast, we did not observe an effect on growth of 4 HL cell lines (L1236, SUPHD1, L428 and L540) after 48, 52 and 68 hours induction with IL-3 by alamar blue assay (Figure 4A). We also didn't see any growth effects after 4 and 24 hours of induction with cytokines (data not shown). In an earlier study by Hsu et al.[21], no effect was observed by exogenous IL-3 on the proliferation of two other HL cell lines (HDLM2 or KMH2). A possible explanation for the discrepancies between these studies may be different culture conditions or different assays.

*IL5RA* mRNA expression was detected in SUPHD1, L1236 and L428 and not in other 4 HL lines (Figure 3D). By flow, we observed weak staining for IL5RA in some of the cell lines, but the signal was close to the level of the isotype control antibody (Figure 3D). *IL-5* mRNA and secreted IL-5 protein (0.69 ng/ml) was detected only in L428 cells (data not shown and Figure 3F), consistent with previous publications[6,7]. In primary HRS cells, IL-5 mRNA has been identified in HRS cells of tumors characterized by presence of tissue eosinophilia.[8] Proliferation of the IL-5 positive HL cell line (L428) was not affected by antibody-mediated neutralization of IL-5, suggesting that IL-5 is not an autocrine growth factor in this cell line.[7] Consistent with these findings, we also did not see an effect on cell growth after the addition of IL-5 (Figure 4B).

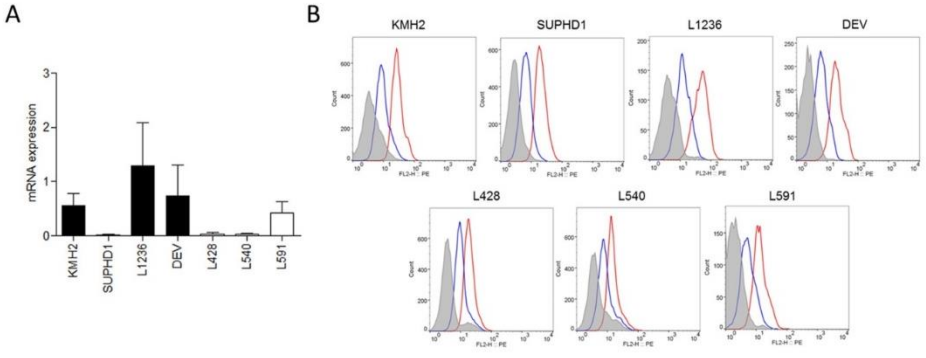



**Figure 2. CSF2RB expression in HL cell lines. (A)** Bar graph showing the qRT-PCR quantification of the relative *CSF2RB* mRNA expression levels (mean ± SD). Mutant cell lines are shown as black bars and the wild type cell lines as white bars; and **(B)** Flow cytometric analysis (mean fluorescence intensity, MFI ± SD) of cell surface expression of CSF2RB protein using PE-labeled mouse anti-CSF2RB antibody (ab93509, Abcam). Gray-filled: blank; Blue line: isotype control; Red line: samples incubated with CSF2RB antibody.



CSF2RA expression levels were very low or undetectable in all seven HL cell lines (data not shown). *GM-CSF* mRNA levels were detected in L1236, L428, L540 and L591 (data not shown). Secreted protein levels in culture supernatant varied between 0.6 and 4.6 ng/ml (Figure 3G). This is consistent with previous studies showing GM-CSF protein expression in two HL cell lines.[6,21] However, in eight primary cHL cases analyzed by Northern analysis, no expression of *GM-CSF* was observed[9] possibly due to the low percentage of HRS cells. Since HL cell lines do not express CSF2RA, they are not dependent on GM-CSF.

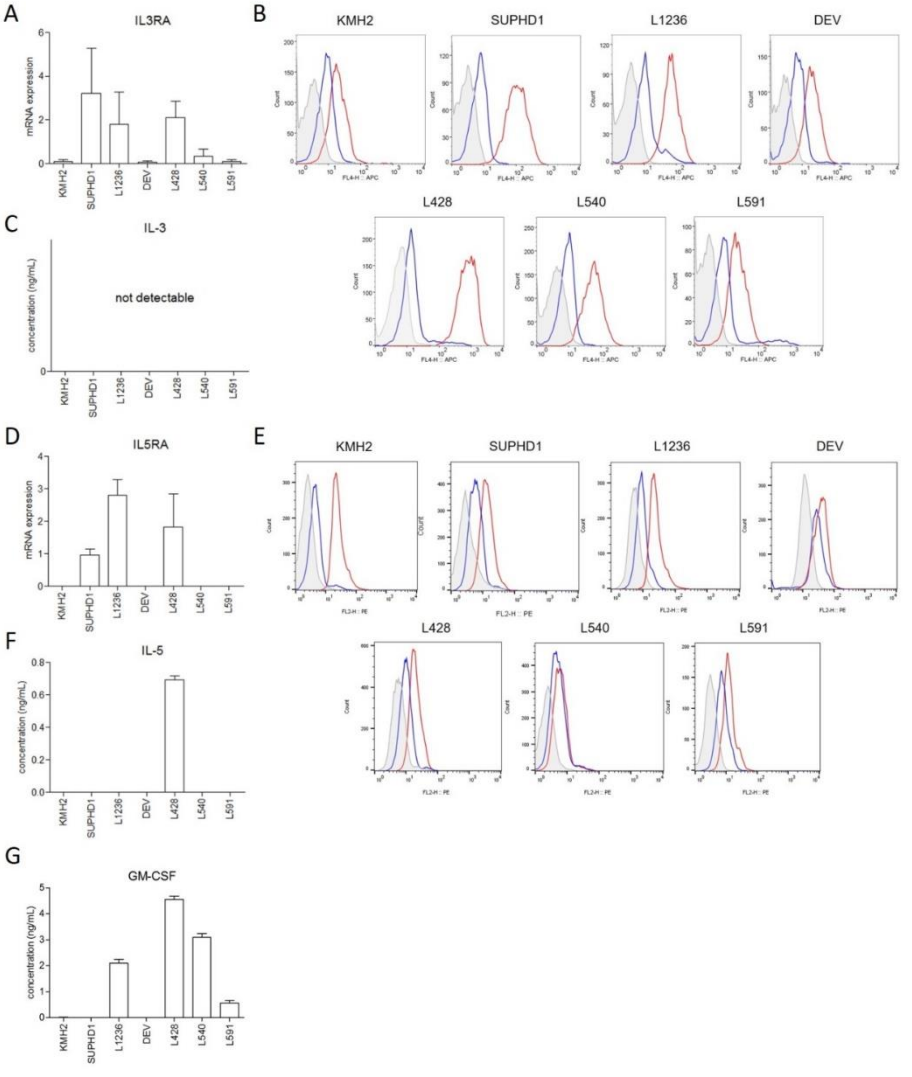



**Figure 3. Expression of the α chain receptors and cytokines in the 7 HL cell lines.** **(A)** Bar graph showing the qRT-PCR quantification of the relative IL3RA mRNA expression (mean ± SD); **(B)** Flow cytometry histogram plots of cell surface IL3RA using APC mouse anti-IL3RA (BD560087, BD Biosciences); **(C)** ELISA analysis of endogenous IL-3; **(D)** Bar graph showing the qRT-PCR quantification of the relative IL5RA mRNA expression levels (mean ± SD); and **(E)** Flow cytometry histogram plots of cell surface IL5RA using rabbit anti-IL5RA (PA5-25159, Thermo Fisher); **(F)** ELISA analysis of endogenous IL-5. **(G)** ELISA analysis of endogenous GM-CSF in 7 HL cell lines. Gray-filled: blank; Blue line: isotype control; Red line: samples incubated with IL3RA or IL5RA antibody.

So far, we did not observe a significant effect on the proliferation of *CSF2RB* mutated (L1236 & SUPHD1) and wild type cell lines (L428 & L540) upon addition of either IL-3 or IL-5. This might be caused by presence of genomic aberrations in genes affecting activation of the JAK/STAT pathway. Activation of the JAK/STAT pathway in HL can be achieved via somatic loss of function mutations of SOCS1[2] and PTPN2[22], two inhibitors of the JAK/STAT pathway. SOCS1 mutations were indeed observed in HDLM2, L1236 and L428 cell lines.[23] HDLM2 cells were shown to have high levels of pSTAT5 possibly due to the SOCS1 mutation. However, pSTAT5 was not detectable in the two other SOCS1 mutated cell lines. This might be due to the nature of the mutations, which were in-frame deletions, resulting in loss of four and five amino acid within the SH2 domain leaving the SOCS box sequence unaffected.[23] In SUPHD1, we found a PTPN2 mutation in our WES data (L249P). This is consistent with a previous study[22] showing that treatment with a JAK kinase inhibitor induced dephosphorylation of all STAT members in SUPHD1 cells and a dose dependent inhibition of the proliferation of the SUPHD1 cells. These findings show that SUPHD1 cells depend on JAK/STAT signaling for their proliferation. In L540, no mutation was found in SOCS1 or PTPN2, but a previous study showed that L540 cells have high pSTAT5 expression.[24] The underlying mechanism is still unknown. JAK2 amplifications have also been observed in 2 of the HL cell lines, i.e. L428 and KMH2. These amplifications result in activation of the JAK/STAT pathway. Furthermore, aberrations in the JAK/STAT pathway have also been reported in primary HL cases.[23,25,26] So it is evident that signaling through JAK/STAT is important for growth of HL cells, but it remains unclear whether and how mutations in CSF2RB might contribute to this.

In summary, we showed *CSF2RB* mutations in four out of seven HL cell lines. With our assays, we did not find an effect of these mutations on growth in HL cell lines. This might be related to somatic mutations or amplification of SOCS1, PTPN2 or JAK2. We did not yet look at the endogenous expression of JAK2, STAT5 and pSTAT5, which might help to design further studies to unravel the role of CSF2RB and the consequences of its mutations in HL.

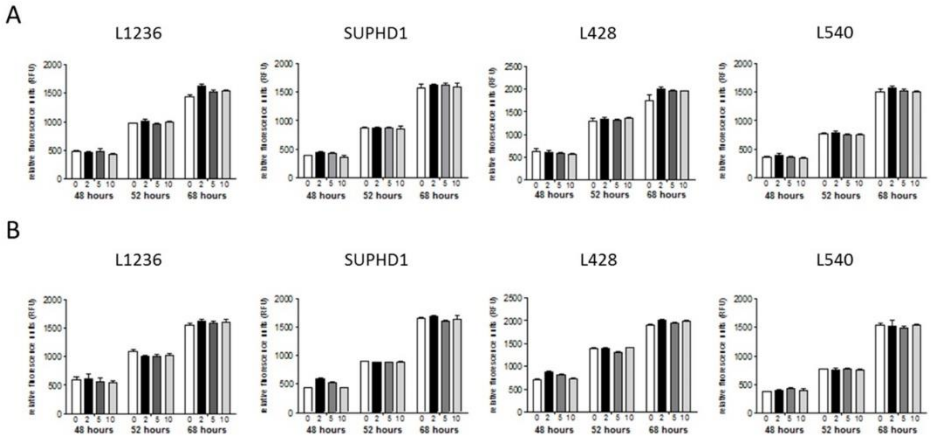



**Figure 4. Alamar blue assay in presence of (A) IL-3 and (B) IL-5 in 4 HL cell lines.** Cell proliferation was measured 48, 52 and 68 hours after adding 0, 2ng/mL, 5ng/mL and 10ng/mL of each of the cytokines. No significant effects were observed in any of the four HL cell lines after introduction of the cytokines.

**Table 1.** Primer sequences for PCR and qPCR.

| Primer | Forward (5' → 3') | Reverse (5' → 3') | Size (bp) |
|---|---|---|---|
| PCR (DNA & cDNA) | | | |
| KMH2_A | GCAGCCATCCTCCTCTCCAA | AGCAAACCTCTGGGCTCCAC | 151 |
| KMH2_B | GGTGGGAGGACAGGACCAAA | CATAGCCCTCAAACCCTGAC | 254 |
| SUPHD1_A (DNA) | CACCCCACCAAGACCCTTGT | AGGCAGCTGGAGTCGTGTCA | 188 |
| SUPHD1_A (cDNA) | CAGCATGTCGGCCTTCACTA | AGGCAGCTGGAGTCGTGTCA | 188 |
| SUPHD1_B | AAGTGGGCGACTATTGCTTCCTC | GGACCGGGTCCCGGGGAAGA | 85 |
| L1236 | CCTGGAGTGGCCTCTGGTTA | CAGGATTGTTCCTTGGTGAC | 241 |
| DEV | CCAGGTCACCAAGGAACAAT | AAGCCTGGTCTAGGTTCTTGAT | 234 |
| qPCR | | | |
| CSF2RB | GTCGTCACTGACGTTGACTA | GTCCTGGTCGGTGCTGATCT | 129 |
| IL3RA | TATCGGGTGTCGTTTCGATG | ACCAGGATGTGGGAACTTTG | 66 |
| IL5RA | CCTGCCTTGTTGGCACAGAT | CTTCAGTCCAAGAGCCATACCT | 75 |
| CSF2RA | TCCTGATCCCAGAGAAATCG | TTACTGAGCCTGGGTTCCAG | 172 |
| IL-3 | TTGCCTTTGCTGGACTTCAA | TTGAATGCCTCCAGGTTTGG | 92 |
| IL-5 | GCTTCTGCATTTGAGTTTGCTAGCT | TGGCCGTCAATGTATTTCTTTATTAAG | 294 |
| GM-CSF | GGAGCATGTGAATGCCATCC | GCAGCAGTGTCTCTACTCAG | 63 |
| TBP | GCCCGAAACGCCGAATAT | CCGTGGTTCGTGGCTCTCT | 73 |

## ACKNOWLEGMENTS

This work was supported by grants from the Abel Tasman Talent Program of the University Medical Center Groningen.

**Table 2.** Aberrations of the JAK/STAT pathway and expression of relevant receptors and cytokines in HL cell lines. Somatic mutations of CSF2RB (c.2198del, V524M, c.2500_2501del, V212I, A715S, Q809*), SOCS1 (c.435_446del, c.191_218del, c.393_407del) and PTPN2 (L249P) in WES[3] data from 7 HL cell lines.

| | L1236 | SUPHD1 | KMH2 | L428 | L540 | L591 | DEV |
|---|---|---|---|---|---|---|---|
| Subtype | mixed cellularity | lymphocyte depleted | mixed cellularity | nodular sclerosis | nodular sclerosis (T cell) | nodular sclerosis | NLPHL |
| Mutation | | | | | | | |
| CSF2RB | c.2198del | V524M & c.2500_2501del | V212I & A715S | no | no | no | Stop gain (Q809*) |
| SOCS1 | c.435_446del & c.191_218del | no | no | c.393_407del | no | no | no |
| PTPN2 | no | L249P | no | no | no | no | no |
| Cell surface receptor expression | | | | | | | |
| CSF2RB | yes | yes | yes | yes | yes | yes | yes |
| IL3RA | low | low | low | high | low | low | low |
| IL5RA | yes | yes | no | yes | no | no | no |
| CSF2RA | no | no | no | no | no | no | no |
| Cytokine level in culture supernatant | | | | | | | |
| IL-3 | no | no | no | no | no | no | no |
| IL-5 | no | no | yes | yes | no | no | no |
| GM-CSF | yes | no | yes | yes | yes | yes | no |
| JAK2 amplification/gain | no[27] | | yes[27] | yes[27] | | | |
| pSTAT5 expression | no[24] | yes[22] | yes[24] | no[24] | yes[24] | | |

## REFERENCES


1. Schmitz R, Stanelle J, Hansmann ML, Kuppers R. Pathogenesis of classical and lymphocyte-predominant Hodgkin lymphoma. Annu Rev Pathol 2009; 4: 151-74.
2. Kuppers R, Engert A, Hansmann ML. Hodgkin lymphoma. J Clin Invest 2012; 122(10): 3439-47.
3. Liu Y, Abdul Razak FR, Terpstra M, Chan FC, Saber A, Nijland M et al. The mutational landscape of Hodgkin lymphoma cell lines determined by whole-exome sequencing. Leukemia 2014; 28(11): 2248-51.
4. Aldinucci D, Poletto D, Gloghini A, Nanni P, Degan M, Perin T et al. Expression of functional interleukin-3 receptors on Hodgkin and Reed-Sternberg cells. Am J Pathol 2002; 160(2): 585-96.
5. Aldinucci D, Olivo K, Lorenzon D, Poletto D, Gloghini A, Carbone A et al. The role of interleukin-3 in classical Hodgkin's disease. Leuk Lymphoma 2005; 46(3): 303-11.
6. Klein S, Jucker M, Diehl V, Tesch H. Production of multiple cytokines by Hodgkin's disease derived cell lines. Hematol Oncol 1992; 10(6): 319-29.
7. Kapp U, Yeh WC, Patterson B, Elia AJ, Kagi D, Ho A et al. Interleukin 13 is secreted by and stimulates the growth of Hodgkin and Reed-Sternberg cells. J Exp Med 1999; 189(12): 1939-46.
8. Samoszuk M, Nansen L. Detection of interleukin-5 messenger RNA in Reed-Sternberg cells of Hodgkin's disease with eosinophilia. Blood 1990; 75(1): 13-6.
9. Merz H, Fliedner A, Orscheschek K, Binder T, Sebald W, Muller-Hermelink HK et al. Cytokine expression in T-cell lymphomas and Hodgkin's disease. Its possible implication in autocrine or paracrine production as a potential basis for neoplastic growth. Am J Pathol 1991; 139(5): 1173-80.
10. Pinto A, Aldinucci D, Gloghini A, Zagonel V, Degan M, Improta S et al. Human eosinophils express functional CD30 ligand and stimulate proliferation of a Hodgkin's disease cell line. Blood 1996; 88(9): 3299-305.
11. Atayar C, Kok K, Kluiver J, Bosga A, van den Berg E, van der Vlies P et al. BCL6 alternative breakpoint region break and homozygous deletion of 17q24 in the nodular lymphocyte predominance type of Hodgkin's lymphoma-derived cell line DEV. Hum Pathol 2006; 37(6): 675-83.
12. Steidl C, Shah SP, Woolcock BW, Rui L, Kawahara M, Farinha P et al. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. Nature 2011; 471(7338): 377-81.
13. Twa DD, Chan FC, Ben-Neriah S, Woolcock BW, Mottok A, Tan KL et al. Genomic rearrangements involving programmed death ligands are recurrent in primary mediastinal large B-cell lymphoma. Blood 2014; 123(13): 2062-5.
14. Reichel J, Chadburn A, Rubinstein PG, Giulino-Roth L, Tam W, Liu Y et al. Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. Blood 2015; 125(7): 1061-72.
15. Freeburn RW, Gale RE, Linch DC. Activating point mutations in the betaC chain of the GM-CSF, IL-3 and IL-5 receptors are not a major contributory factor in the pathogenesis of acute myeloid leukaemia. Br J Haematol 1998; 103(1): 66-71.
16. Watanabe-Smith K, Tognon C, Tyner JW, Meijerink JP, Druker BJ, Agarwal A. Discovery and functional characterization of a germline, CSF2RB-activating mutation in leukemia. Leukemia 2016.
17. Tanaka T, Motoi N, Tsuchihashi Y, Tazawa R, Kaneko C, Nei T et al. Adult-onset hereditary pulmonary alveolar proteinosis caused by a single-base deletion in CSF2RB. J Med Genet 2011; 48(3): 205-9.

**18.** Jenkins BJ, D'Andrea R, Gonda TJ. Activating point mutations in the common beta subunit of the human GM-CSF, IL-3 and IL-5 receptors suggest the involvement of beta subunit dimerization and cell type-specific molecules in signalling. Embo J 1995; 14(17): 4276-87.
**19.** Jenkins BJ, Blake TJ, Gonda TJ. Saturation mutagenesis of the beta subunit of the human granulocyte-macrophage colony-stimulating factor receptor shows clustering of constitutive mutations, activation of ERK MAP kinase and STAT pathways, and differential beta subunit tyrosine phosphorylation. Blood 1998; 92(6): 1989-2002.
**20.** Bosshart H. Interleukin-3 receptors in Hodgkin's disease. Am J Pathol 2003; 162(1): 355-6; author reply 356-7.
**21.** Hsu SM, Hsu PL. Lack of effect of colony-stimulating factors, interleukins, interferons, and tumor necrosis factor on the growth and differentiation of cultured Reed-Sternberg cells. Comparison with effects of phorbol ester and retinoic acid. Am J Pathol 1990; 136(1): 181-9.
**22.** Kleppe M, Tousseyn T, Geissinger E, Kalender Atak Z, Aerts S, Rosenwald A et al. Mutation analysis of the tyrosine phosphatase PTPN2 in Hodgkin's lymphoma and T-cell non-Hodgkin's lymphoma. Haematologica 2011; 96(11): 1723-7.
**23.** Weniger MA, Melzner I, Menz CK, Wegener S, Bucur AJ, Dorsch K et al. Mutations of the tumor suppressor gene SOCS-1 in classical Hodgkin lymphoma are frequent and associated with nuclear phospho-STAT5 accumulation. Oncogene 2006; 25(18): 2679-84.
**24.** Skinnider BF, Elia AJ, Gascoyne RD, Patterson B, Trumper L, Kapp U et al. Signal transducer and activator of transcription 6 is frequently activated in Hodgkin and Reed-Sternberg cells of Hodgkin lymphoma. Blood 2002; 99(2): 618-26.
**25.** Martini M, Hohaus S, Petrucci G, Cenci T, Pierconti F, Massini G et al. Phosphorylated STAT5 represents a new possible prognostic marker in Hodgkin lymphoma. Am J Clin Pathol 2008; 129(3): 472-7.
**26.** Van Roosbroeck K, Cox L, Tousseyn T, Lahortiga I, Gielen O, Cauwelier B et al. JAK2 rearrangements, including the novel SEC31A-JAK2 fusion, are recurrent in classical Hodgkin lymphoma. Blood 2011; 117(15): 4056-64.
**27.** Joos S, Granzow M, Holtgreve-Grez H, Siebert R, Harder L, Martin-Subero JI et al. Hodgkin's lymphoma cell lines are characterized by frequent aberrations on chromosomes 2p and 9p including REL and JAK2. Int J Cancer 2003; 103(4): 489-95.

# Chapter 5

## *MYB* gene mutations and oncogenic properties in Hodgkin lymphoma

**Fazlyn Reeny Abdul Razak**, Arjan Diepstra, Michelle Oomkens, Debora de Jong, Jasper Koerts, Bea Rutgers, Lydia Visser, Joost Kluiver, Anke van den Berg

Manuscript in preparation

**ABSTRACT**

Hodgkin lymphoma (HL) is characterized by constitutive activation of several signaling pathways and expression of a distinct set of transcription factors. This characteristic phenotype is partly caused by gene mutations as determined by targeted and whole genome sequencing approaches of cell lines and microdissected HRS cells. In our WES data, we found frame shift mutations in *MYB* in 2 out of 7 HL cell lines. The MYB transcriptional activator is involved in the regulation of proliferation and differentiation of hematopoietic cells. The frame shift mutations in L428 and SUPHD1 were confirmed at the DNA and RNA level by Sanger sequencing. Presence of truncated protein was confirmed by Western blot in these two cell lines and expression of full length MYB protein was shown in the other HL cell lines. In contrast to the cell lines, MYB staining was negative in almost all tested primary diagnostic HL cases. Inhibition of MYB using shRNA constructs induced negative effects on cell growth in two HL cell lines with wild type MYB (L540 & KMH2) and in one cell line with truncated MYB protein (SUPHD1). There was no difference in cell cycle distribution, but a marked induction of apoptosis was revealed. MYB protein levels are regulated by MYC, whereas MYC levels are independent of MYB in HL. Microarray analyses of L540 & KMH2 cells infected with non-targeting and MYB shRNA constructs revealed 60 consistently MYB regulated genes, including MYB induced anti-apoptotic BCL2. Gene ontology analysis indicated that MYB gene functions are related to proliferation, apoptosis, cell adhesion and immune response. In conclusion, we found strong oncogenic effects of wild type and truncated MYB in HL cell lines. These effects on cell growth were most likely caused by repression of apoptosis, probably mediated by BCL2.

## INTRODUCTION

Hodgkin lymphoma (HL) accounts for 11% of all malignant lymphomas with around 66,000 new cases being diagnosed in Europe in 2012.[1] HL is characterized by a minority of B-cell derived tumor cells, named Hodgkin Reed-Sternberg (HRS) cells in classical (c)HL and lymphocyte predominant (LP) cells in nodular lymphocyte predominant (NLP) HL. HRS cells lack expression of the B-cell receptor in most cases and rely on activation of multiple pathways to escape from apoptosis in the germinal center reaction. Our previously reported whole exome sequencing (WES) analysis of seven HL cell lines revealed 373 recurrent mutations.[2] One of the recurrent mutations involved the MYB gene, which was mutated in 2 HL cell lines. MYB proteins are a family of transcription factors that are involved in the regulation of differentiation and proliferation.[3] MYB was originally identified as the cellular counterpart of the v-Myb oncogene found in avian leukemia viruses. These viruses transform immature hematopoietic cells *in vitro* and induce acute leukemias in chickens.[4-6] The MYB protein has a highly conserved DNA binding domain (DBD) near its N-terminus and a large transcriptional activation domain (TAD) and a negative regulation domain (NRD) in the C-terminal region.[7-9] The NRD contains multiple subdomains, including a putative leucine zipper motif/heptad leucine repeat. Although the mechanisms of negative regulation by NRD are not fully understood, it is likely to act through intramolecular interactions,[8] interaction with other proteins,[10-12] and effects on protein stability[13]. Thus, the large C-terminal domain, including both TAD and NRD, has multiple functional components involved in regulating both the specificity and the activity of MYB.[14]



Deregulated expression of MYB has been implicated in many human malignancies, including several types of leukemias, carcinomas (colon, breast, pancreas and head and neck), melanomas and glioblastomas.[15-17] Previous studies suggested that loss of the NRD region can be a driver of oncogenesis. One of the mechanisms affecting the NRD is alternative RNA splicing.[3,18] The *MYB* gene exhibits a quite complicated alternative splicing pattern resulting in transcripts that encode for truncated MYB protein variants. These variants include the complete DBD but have lost a part of the TAD and the entire NRD. Alternatively spliced *MYB* transcripts have been detected in normal and tumor cells.[18-21] Elevated levels of alternatively spliced *MYB* transcripts have been shown in leukemia samples.[18,20-23] The alternatively spliced exon 9A transcript variant produces a protein that is truncated and resembles the Avian myeloblastosis virus (AMV) v-Myb protein suggesting a putative oncogenic role. Another mechanism affecting the C-terminal domain of MYB is a translocation between chromosomes 6 and 9, resulting in a MYB / NFIB fusion product.[17,24] Moreover, amino acid substitutions in TAD have been implicated in

the transforming activity of v-Myb.[25-30] Although these data do support a role of truncated MYB, there is no direct functional evidence supporting an oncogenic role of alternatively spliced or mutant MYB.[14]

A substantial number of MYB target genes have been identified including genes that have diverse functions in processes such as myeloid differentiation, cellular proliferation, cell cycle, apoptosis and cell signaling.[31] BCL2 has been shown to be a direct MYB target.[32-34] Several studies indicated that *MYC* might be a potential target gene of MYB based on the presence of MYB binding sites in the promotor region of MYC.[35,36] Vice versa, some studies also suggested that MYC indirectly targets the MYB expression through microRNA-150 (mir-150).[37]

Thus, it has become clear that the functionality of MYB is subject to change through mechanisms that alter its C-terminal domain such as alternative splicing, translocations, and mutations. In this study, we validated *MYB* gene mutations and determined the role of wild type and mutant *MYB* in HL cell lines by determining effects on tumor cell growth and by defining the set of MYB-regulated genes.

## MATERIALS AND METHODS

### *Cell lines and primary samples*

HL cell lines DEV (nodular lymphocyte predominant HL, cell line developed in house),[38] L428 (nodular sclerosis, DSMZ No ACC 197), L1236 (mixed cellularity, DSMZ No ACC 530), KMH2 (mixed cellularity, DSMZ No ACC 8), L591 (nodular sclerosis, EBV+, DSMZ No ACC 602) and L540 (nodular sclerosis, T cell derived, DSMZ No ACC 72) were cultured in RPMI 1640 medium (Lonza Walkersville, Walkersville, MD) supplemented with 5% (L428), 20% (DEV, L540) and 10% (other cell lines) fetal calf serum, 100U/ml penicillin/streptomycin and ultraglutamine (Lonza Walkersville) in a 5% CO2 atmosphere at 37°C. SUPHD1 (lymphocyte depleted subtype, DSMZ No ACC 574) was cultured in McCoy 5A medium supplemented with 10% fetal calf serum. The origin of all cell lines was confirmed with STR DNA analysis 2 to 3 times per year by using Promega Powerplex 16 HS system (DC2101) (Promega, Madison, USA). Mycoplasma tests were performed routinely and showed that the cell lines were not contaminated. Primary diagnostic tissue samples of 37 HL patients were retrieved from the tissue bank of the Pathology Department, University Medical Center Groningen. The study was conducted in

accordance with the Declaration of Helsinki and the national Code of conduct for responsible use of human tissue (FEDERA) guidelines.

***DNA and RNA isolation***

DNA was isolated using the salt-chloroform extraction method and RNA was isolated using the miRNeasy micro kit (Qiagen, Venlo, The Netherlands) following the manufacturer's instructions. Micro Bio-Spin chromatography columns supplied with Bio-Gel P-6 polyacrylamide gel matrices, were used to purify RNA samples (Bio-Rad laboratories B.V. Veenendaal, The Netherlands). DNA and RNA concentrations and purity were measured on a Nanodrop ND-1000 Spectrophotometer (Nanodrop Technologies, Wilmington, USA). Only good quality DNA and RNA samples were used for further analysis.

***Sanger sequencing***



M13-tail containing primers were designed to amplify the region containing the known *MYB* mutations at the DNA and mRNA level (Table S1). Amplification reactions were performed following standard protocols provided by the manufacturer. PCR products were gel purified using high pure PCR product purification kit (Roche, Mannheim, Germany) and send for Sanger sequencing (LGC Genomic, Teddington, UK).

***cDNA synthesis, and (Qualitative) RT-PCR***

Complementary DNA (cDNA) was synthesized from 500ng of total RNA using Superscript II RT (Invitrogen, Carlsbad, USA) and random hexamers in a final volume of 20µl. PCR was performed in triplicate using 10ng cDNA and 500nM primers in a final volume of 30µl according to the manufacturer's instructions (Invitrogen) for 35 cycles using a thermocycler (Bio-Rad). Appropriate positive and negative controls were included in each step. PCR products were analyzed on a 1% agarose gel. Quantitative (q)RT-PCR was performed in triplicate in a volume of 10µl using SYBR Green master mix and the Lightcycler 480 (Roche). TBP was used as a housekeeping gene for normalization. Relative expression levels are expressed as $2^{-(\text{delta Cp})}$. Primers used for the amplification are listed in Table S1.

To determine presence of alternatively spliced MYB transcripts for the smaller exons, we used fluorescently labeled primers in the amplification. The amplified fragments were separated via capillary electrophoresis using Applied Biosystems 3500 Genetic Analyzer. The capillary electrophoresis (CE) was performed with 8.7µL HiDi Formamide (4311320) (Applied Biosystems, Foster City, CA) and 0.3µL Gene Scan600 LIZ Size Standard (4408399) (Applied

Biosystems) and 1µL PCR product. The samples were set up in a 96 well PCR plate, heated for 3 min at 95C° and cooled on ice for 3 minutes. The data were analyzed using GeneMapper software (version 3.7, Applied Biosystems).

***Western blotting***

Cells were lysed in Radio Immuno Precipitation Assay (RIPA) lysis buffer (Millipore, Amsterdam, The Netherlands). Protein concentration was measured with the Pierce BCA Protein Assay Kit (Thermo Scientific, Massachusetts, USA). 20µg protein was loaded onto SDS-polyacrylamide electrophoresis gel and followed by Western blot with anti-MYB monoclonal antibody, N-terminus (1:500, ab45150, Abcam, Cambridge, UK) and C-terminus (1:500, ab169111, Abcam). As a loading control, we analyzed expression of GAPDH (1:20000, Clone AC-15, Sigma, St. Louis MO). Staining was visualized using Supersignal ® Chemiluminescent Substrate (Thermo Scientific). Detection and quantification of band intensities was conducted using Image Lab software (version 5.2, Bio-Rad).

***Immunohistochemistry***

Immunostaining was performed with two anti-MYB monoclonal antibodies, one binding to the N-terminus (1:100, ab45150, Abcam) and one binding to the C-terminus (1:40, ab169111, Abcam). Paraffin embedded HL cell lines and primary diagnostic HL tissue sections were subjected to antigen retrieval in EDTA buffer (pH 8.0) in a pressure cooker for 15 minutes. Tissue sections were incubated with primary antibody at 4°C overnight. Positive staining was visualized after secondary antibody incubation steps with 3,3'-Diaminobenzidine. The tissues were examined and scored by an experienced hematopathologist.

***shRNA knockdown***

Custom designed shRNAs (Table S1) were cloned into the pGreenPuro lentiviral vector (System Biosciences, Mountain View, USA) that in addition to the shRNA cassette also contained a GFP gene. Two shRNA constructs against human MYB (MYB#1 and MYB#2), two against MYC (MYC#1 and MYC#2) and two non-targeting control (NT1 and NT2) were generated. Lentiviral particles were produced by calcium phosphate mediated transfection of HEK293T cells with 2.5M $CaCl_2$, 1µg pMSCV-VSV-G, 1µg pRSV.REV and 1µg pMDL-gPRRE. Supernatant was collected after 48 hours. Lentiviral particles were passed through a 0.45µM millex-HV PVDF filter (Millipore) before infection. HL cells (L540, KMH2, L428 and SUPHD1) were transduced in the presence of 4µg/mL of polybrene (Sigma). GFP+ cells were sorted using

the MoFlo flow cytometry cell sorter (Beckman Coulter) 7 days after infection to check efficiency of the shRNA constructs. The efficiency was tested at the RNA level by qRT-PCR and at the protein level by Western blot. The knockdown efficiency was calculated using the formula: (normalized MYB or MYC expression of MYB or MYC shRNA sample/ normalized MYB or MYC expression of NT sample) x 100. Functional consequences on proliferation, apoptosis and cell cycle distribution were measured on day 9 as indicated below.

***GFP competition assay***

HL cell lines (L540, KMH2, L428 and SUPHD1) were infected with MYB, MYC and control shRNA constructs aiming at an infection rate between 20-45%. The percentage of GFP+ cells was analyzed triweekly for 22 days. Data were acquired on a FACS Calibur flow cytometer (BD PharMingen, San Diego, USA) and analyzed using FlowJo (version 7.6, Treestar, Ashland, OR). The GFP percentages were normalized to the GFP percentage at day 4 after infection. The GFP competition assay was performed in triplicate for each of the cell lines. The average decline of GFP in the experiments was used to determine whether the decline in GFP percentage of MYB knockdown cells was significantly different from the GFP percentages of the cells infected with control.



***Cell cycle Analysis***

Sorted GFP+ cells were washed with cold PBS and incubated with PI/Triton X-100 staining solution. Cells were lysed on ice for approximately 10 min before measuring by FACS Calibur flow cytometry (BD PharMingen). Doublet discrimination was applied to remove cell clumps and enable the analysis of single cell events by using PI area versus PI width signal. Cells were analyzed for DNA content and the percentage of cells in each phase of the cell cycle (G1, S, and G2/M) was calculated using ModFit LT (version 4.1.7, Verity Software House, Topsham ME).

***Apoptosis assay***

Sorted GFP+ cells were incubated with annexin V-APC (BD Biosciences, New Jersey, USA) in calcium buffer for 15 minutes in the dark. Stained cells were immediately subjected to flow cytometry analyses using FACS Calibur flow cytometer (BD PharMingen) and analyzed using FlowJo software (version 7.6, Treestar, Ashland, OR). Based on the staining pattern, cells were grouped as live cells (annexin V-) or apoptotic cells (annexin V+).

### *Microarray analysis*

Microarray experiments were performed on a custom designed microarray based on the SurePrint G3 8x60K array format that contained probes for all protein-coding genes (30,094 probes) (Agilent Technologies, Amstelveen, NL).[39] The samples were Cy3 labeled using 50-100ng total RNA. After labeling samples were purified using the RNeasy Mini Kit (74106, Qiagen). For hybridization, we used the Gene Expression Hybridization Kit (Agilent Technologies). Arrays were scanned with the Agilent DNA Microarray Scanner and data was generated with Agilent Feature Extraction software. Resulting raw data were analyzed using Gene Spring GX software version 12.5 (Agilent Technologies). Quantile normalization was used for all samples. Probe sets were filtered based on "flagged present". Significant expression changes in MYB shRNA and control shRNA infected cells were determined by a moderated Student's t test. Heatmaps were generated in Genesis (Genomics and Bioinformatics Graz, Graz, Austria) using unsupervised hierarchical clustering with Pearson correlation and median centering of the data.

### *Gene Ontology (GO) Analysis*

Gene Ontology enrichment analysis of was performed using the DAVID Gene Functional Classification Tool and g-Profiler Gene Ontology enrichment analysis tool.

### *Statistical analysis*

Results obtained by qRT-PCR are expressed as mean ± standard deviation respectively. For the GFP competition assay, the average decline in proliferation was determined. A p-value $<0.05$ was considered significant. Statistical analyses were performed with GraphPad Prism version 5.0 (GraphPad Software, San Diego, CA, USA) and SPSS statistical software (IBM SPSS Statistics 20; IBM Corporation, NY, USA).

## RESULTS

### *MYB* mutations and expression in HL cell lines

In our previous whole exome sequencing study, we reported *MYB* frame shift mutations in 2 out of 7 HL cell lines, i.e. L428 and SUPHD1. We now confirmed these two mutations in L428 and SUPHD1 at the DNA and mRNA levels by Sanger sequencing (Figure 1A). Consistent with the exome-seq data, we only observed mutant and no wild type sequences in the two cell lines. The mutations resulted in truncated MYB protein variants resembling alternatively spliced MYB transcripts 8A (L428) and 10A (SUPHD1) (Figure 1B).

There were no differences at the *MYB* mRNA levels between the mutated cell lines (L428 and SUPHD1) and the non-mutated HL cell lines (Figure 1C). Analysis of alternative splicing in all 7 HL cell lines with primers flanking each known alternatively spliced exon indicated that there were no alternatively spliced transcripts detectable in the HL cell lines (Figure 1D). For the splice variants differing less than 10bp in size, we confirmed this by fragment analysis (data not shown).

Western blotting using a monoclonal antibody binding to the N-terminus of the MYB protein revealed a ~72 kDa band corresponding to the full-length MYB protein in KMH2, L540, L1236, L591 and DEV. A band corresponding to the truncated MYB protein was detected in L428 (~41kDa) and in SUPHD1 (~51 kDa). A monoclonal antibody directed against the C-terminus revealed a band in the five wild type cell lines and no band in L428 and SUPHD1 consistent with the observed mutations (Figure 1C). The bands around 72 kDa in SUPHD1 observed for the N-terminus antibody are not visible with the C-terminus antibody, indicating that these signals are aspecific. Staining of MYB in primary HL tissue sections revealed no staining with both antibodies in the vast majority of the cases (data not shown), which might indicate that gain of MYB expression is a late event in HL.

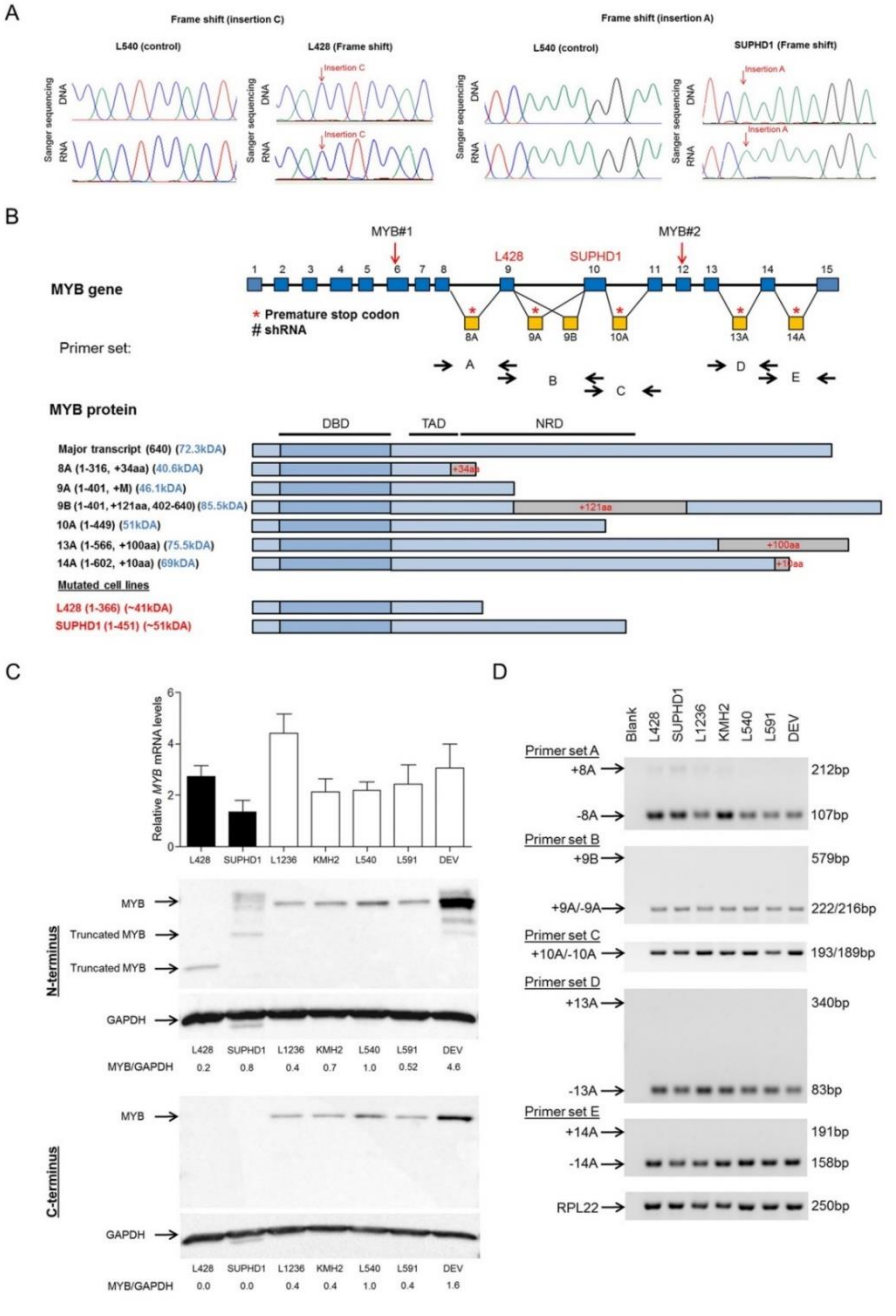

A
Frame shift (insertion C)
L540 (control)
L428 (Frame shift)
Frame shift (insertion A)
L540 (control)
SUPHD1 (Frame shift)
Sanger sequencing
DNA
RNA
Insertion C
Insertion A
B
MYB gene
MYB#1
MYB#2
L428
SUPHD1
* Premature stop codon
# shRNA
Primer set:
MYB protein
DBD
TAD
NRD
Major transcript (640) (72.3kDA)
8A (1-316, +34aa) (40.6kDA)
9A (1-401, +M) (46.1kDA)
9B (1-401, +121aa, 402-640) (85.5kDA)
10A (1-449) (51kDA)
13A (1-566, +100aa) (75.5kDA)
14A (1-602, +10aa) (69kDA)
Mutated cell lines
L428 (1-366) (~41kDA)
SUPHD1 (1-451) (~51kDA)
C
Relative MYB mRNA levels
L428
SUPHD1
L1236
KMH2
L540
L591
DEV
N-terminus
MYB
Truncated MYB
GAPDH
MYB/GAPDH
0.2
0.8
0.4
0.7
1.0
0.52
4.6
C-terminus
0.0
0.0
0.4
0.4
1.0
0.4
1.6
D
Blank
Primer set A
+8A
-8A
212bp
107bp
Primer set B
+9B
+9A/-9A
579bp
222/216bp
Primer set C
+10A/-10A
193/189bp
Primer set D
+13A
-13A
340bp
83bp
Primer set E
+14A
-14A
191bp
158bp
RPL22
250bp

**Figure 1. MYB in HL cell lines. (A)** Validation by Sanger-seq at the DNA and RNA level, showing frame shift insertions in L428 and SUPHD1. **(B)** Schematic diagram of the *MYB* gene and MYB proteins. The black arrows indicate the primer position and red arrows indicate the shRNA position. Red words indicate the position of *MYB* mutations in L428 and SUPHD1 by WES. **(C)** *MYB* levels in HL by qRT-PCR (mean with SD) and western blot (N-terminus and C-terminus). Black arrow indicate the truncated protein in L428 and SUPHD1. The bands at 72 kDa in SUPHD1 with N-terminus antibody, are probably aspecific bands, as no bands are observed with the C-terminus antibody. **(D)** RT-PCR for detection of alternative exons. For primer sets B and C, a fragment analysis was done on the extracted cDNA to allow discrimination between normal and alternative spliced products. Only the normal transcript was detected.

### *Consequences of MYB downregulation in HL cells*

To silence the *MYB* gene, we infected the L540 and KMH2, MYB wild type cell lines and the L428 and SUPHD1, MYB mutant cell lines with MYB shRNA. The efficiency of the MYB knockdown was 13% to 75% at the mRNA level and 37% to 93% at the protein level. The inhibition of MYB expression was more pronounced in the two cell lines with wild type MYB and less evident in the two cell lines with mutant MYB (Representative images of two cell lines are shown in Figure 2A and 2B).



As endogenous MYC levels are high in HL cell lines (Figure S2), we also tested a possible regulation of MYC by MYB and vice versa. MYB knockdown did not alter MYC mRNA levels or protein levels (Figure 2C and 2D). The efficiency of MYC knockdown in the four HL cell lines was limited, 15% to 45% at the mRNA, but more pronounced, 20% to 75%, at the protein level (Figure 3A and 3B). After knockdown of MYC we observed a marked reduction in MYB protein levels in L540 cells with wild type MYB and a limited reduction in MYB protein in L428 cells carrying mutant MYB (Figure 3C and 3D). Thus, we observed a reduced MYB protein expression level after knockdown of MYC in HL cells. We also checked the expression of miR-150 in our knockdown samples as MYC might regulate MYB through induction of miR-150. However, the levels of miR-150 were very low and undetectable in all of our samples including MYB and MYC shRNA treated cells (data not shown).

Inhibition of MYB induced negative effects on cell growth in two HL cell lines with wild type MYB (L540 & KMH2) and in one cell line with truncated MYB (SUPHD1) ($p<0.01$ in all three cell lines). No effect was observed in the second HL cell line with a truncating MYB mutation (L428) (Figure 4A). This might be consistent with the lower effectiveness of the shRNAs as shown on the mRNA and protein levels (Figure 2B). However, it is unclear why the effectiveness of

the shRNA constructs differ between the cell lines, as endogenous MYB levels are quite similar (Figure 1C). To investigate the cause of the growth inhibitory effect, we analyzed cell cycle distribution and induction of apoptosis in the HL cells at day 9 after infection. No difference was observed in the cell cycle distribution upon MYB knockdown in any of the 4 HL cell lines (Figure 4B). An increased number of apoptotic cells was observed in MYB shRNA samples compared to the control in 2 HL cell lines (Figure 4C).

To explore an effect of MYC knockdown on growth of HL cells, we performed a GFP competition assay in L540, KMH2, L428 and SUPHD1 cells. Knockdown of MYC had a strong negative effect on the proliferation of L540 (64% in three weeks) and SUPHD1 (76% in three weeks) cells ($p<0.001$) with both shRNA constructs. In KMH2, a moderate suppression ($p<0.05$) of cell growth was seen (27% in three weeks), whereas only limited effects were observed in the L428 cells (Figure 5A). Cell cycle and an apoptosis analysis were done to further investigate the growth inhibitory mechanism of MYC knockdown in HL cells. No evident difference was observed in cell cycle distribution in any of the HL cell lines (Figure 5B). An increased number of apoptotic cells was observed in the two HL cell lines with the most pronounced negative effect of MYC knockdown on cell growth. No difference was observed in the percentage of apoptotic cells in KMH2 and L428 cells (Figure 5C).

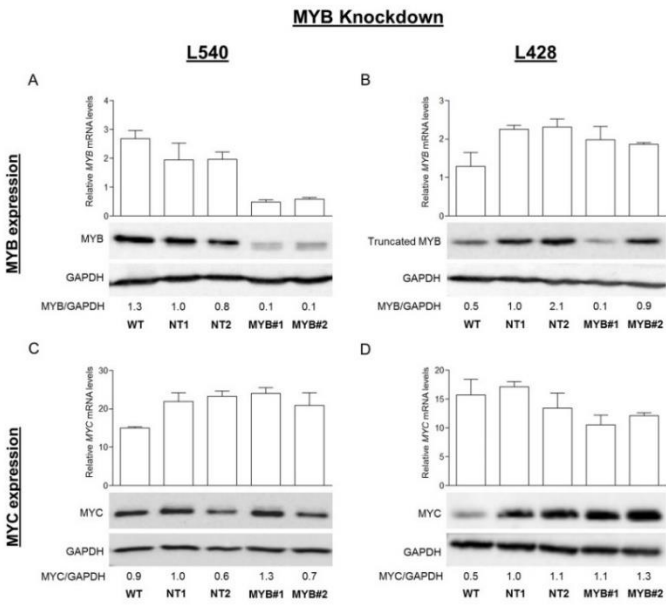



**Figure 2.**
**Two representative examples of the MYB knockdown efficiency by shRNA.** MYB levels were measured at the mRNA (qRT-PCR) and protein level (western blot) levels in a **(A)** MYB WT cell line (L540) and **(B)** MYB mutant cell line (L428). MYC expression was determined by qRT-PCR and western blot in MYB shRNA treated samples in **(C)** L540 and **(D)** L428. RNA isolation and protein lysates were prepared 7 days after transfections. mRNA levels were normalized to TBP.



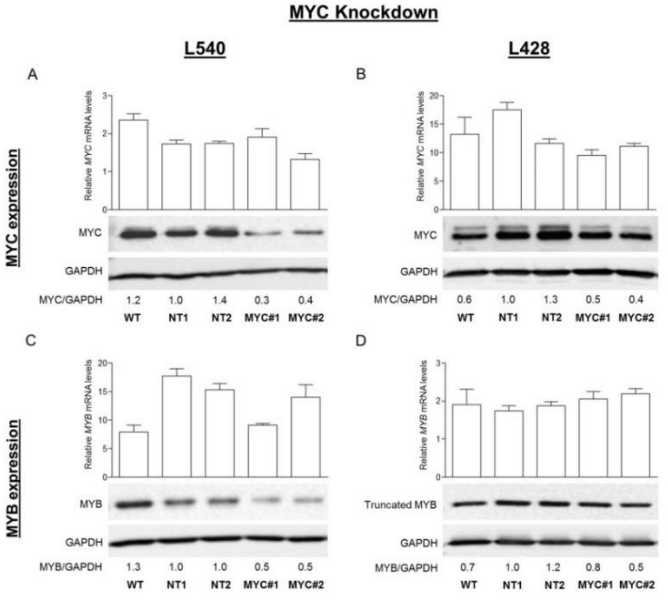



**Figure 3.**
**MYC knockdown efficiency by shRNA in two representative HL cell lines (A)** L540 and **(B)** L428. MYC levels were measured at the mRNA (qRT-PCR) and protein level (western blot) 7 days after transfections. MYB expression was determined upon MYC knockdown by qRT-PCR and western blot in the same two cell lines **(C)** L540 and **(D)** L428.

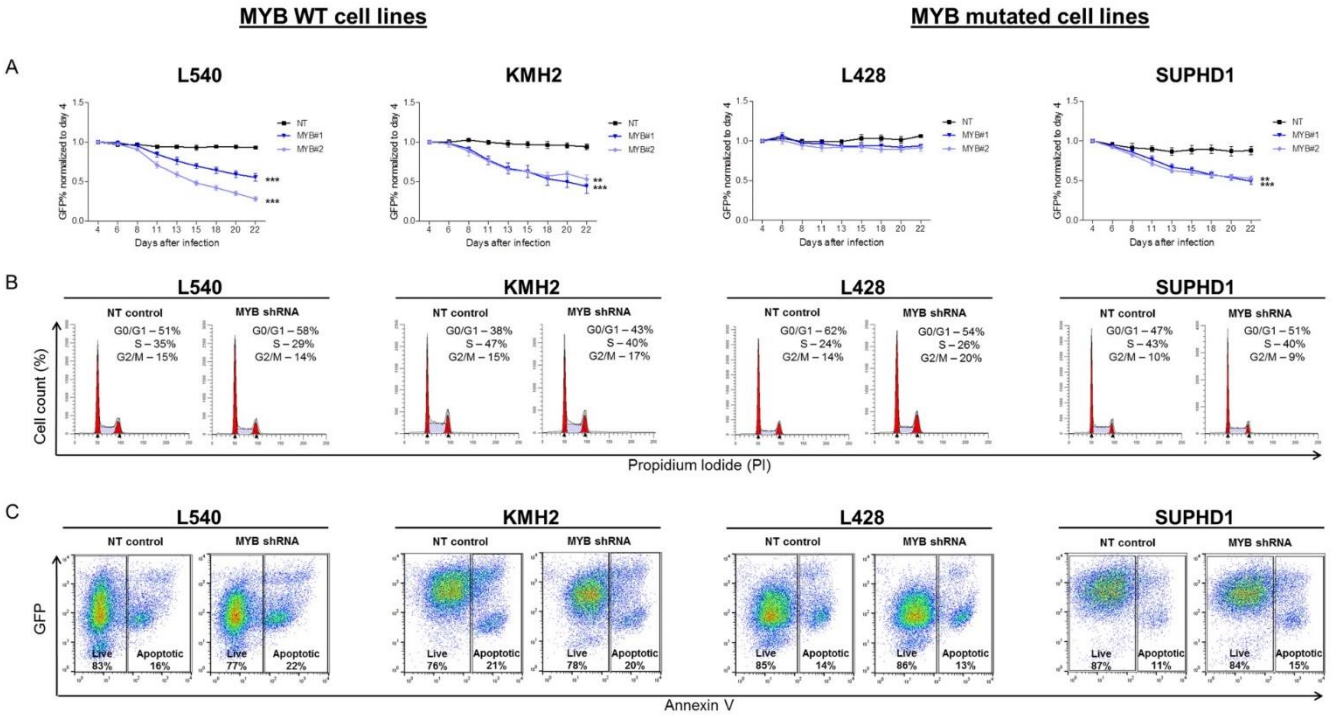



**Figure 4. Effect of MYB knockdown on proliferation, cell cycle and apoptosis in HL cells**. **(A)** Inhibition of MYB resulted in negative effects on cell growth in two HL cell lines expressing wild type MYB (L540 and KMH2) ) and in one cell line with truncated MYB (SUPHD1). No effect was observed in the second HL cell line with a truncating MYB mutation (L428). Shown is the normalized change in percentage of GFP+ cells compared to day 4. **(B)** DNA histograms of HL cells using flow cytometry with propidium iodide (PI) staining revealed no changes in cell cycle distribution. **(C)** Dot-blot graphs of annexin V stained HL cells to assess apoptosis using flow cytometry revealed an increase in the percentage of apoptotic cells in two of the four cell lines. At day 9 after infection, cells were stained with annexin V and PI then analyzed for DNA content by FACS. HL cell lines were transfected with NT and two MYB shRNA constructs, one representative MYB shRNA sample is shown. ***p<0.001, **p<0.01

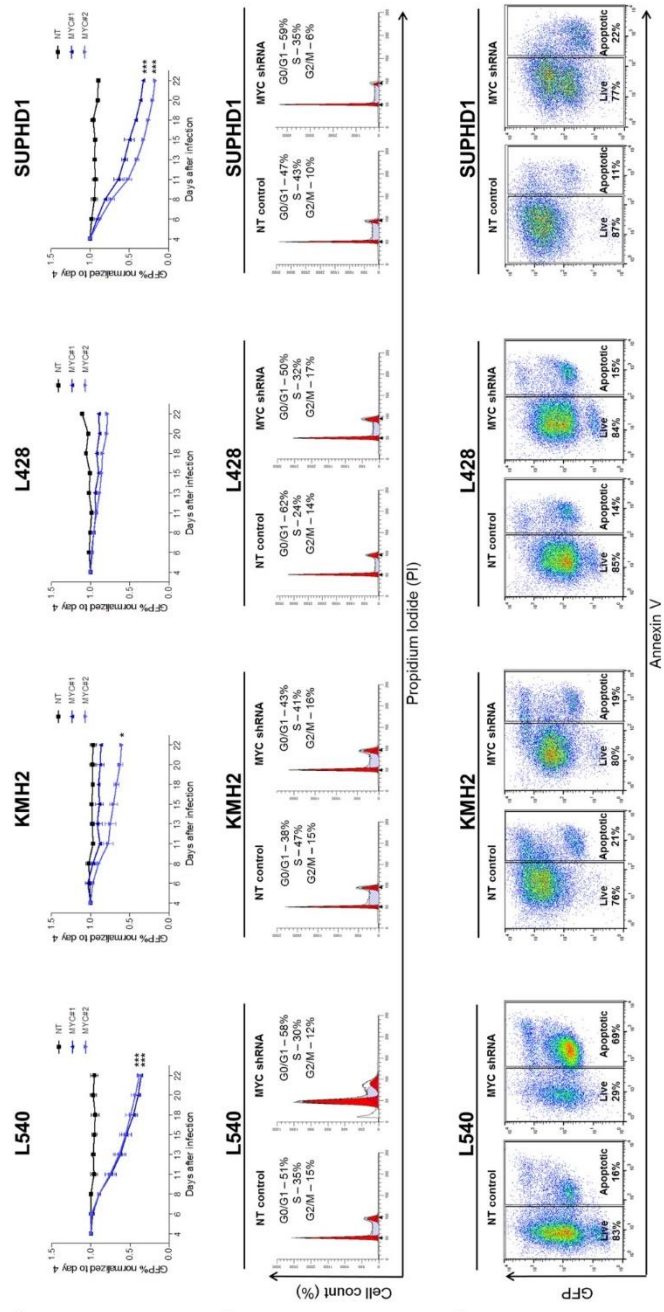



**Figure 5. Effect of MYC knockdown on proliferation, cell cycle and apoptosis in HL cells**. **(A)** Knockdown of MYC had a strong negative effect on the proliferation of L540 and SUPHD1 cells with both shRNA constructs. In KMH2, a moderate suppression of cell growth was seen whereas only limited effects were observed in the L428 cells. The percentage of GFP+ cells was followed over time to establish the effect on cell growth. Shown is the normalized change in percentage of GFP+ cells on day 22 post transduction compared to day 4. At day 9 after infection, cells were stained with Annexin V and PI then analyzed for DNA content by FACS. **(B)** Cell cycle analysis of HL cells using flow cytometry with propidium iodide (PI) staining and the DNA histograms. **(C)** Assessment of apoptosis using flow cytometry with annexin V staining and the dot-plot graph. Representative samples for NT control and MYC shRNA are shown. ***p<0.001. *p<0.05

***Target genes of MYB in L540 and KMH2 cells***

To identify downstream target genes of MYB that could be involved in the growth inhibitory effects in L540 and KMH2 cells, we performed microarray analysis of GFP positive cells at day 7. Filtering by flags resulted in 15,840 probes in L540 and 15,971 probes in KMH2. A moderate T-test ($p<0.05$) (without multiple testing correction) revealed 1,474 differentially expressed probes (1,409 genes) in L540 with a fold change varying between 1.2 and 11.5 for up- and between 1.2 and 21.2 for the down-regulated probes. For KMH2 we found 1,481 differentially expressed probes (1,426 genes) with a fold change varying between 1.2 and 9.4 for the up- and between 1.2 and 4.9 for the down-regulated probes (Figure S1). Probes corresponding to 60 protein coding-genes were consistently differentially expressed in both cell lines (35 down- and 25 up-regulated) (Figure 6A and 6B and Table 1). As expected, MYB was consistently down-regulated in both HL cell lines.

Gene ontology analysis of the molecular function, pathway and biological process was performed to identify genes features of the 60 consistent genes. This revealed three genes related to proliferation (TP53I11, KIFAP3, STAT1), two related to apoptosis (BCL2, WWOX), three related to cell adhesion (PTPN22, ABAT, CD247) and two related to immune response (CIITA, TLR5). As BCL2 has been shown to be a direct MYB target, we validated BCL2 downregulation in MYB knockdown HL cells by western blot (Figure 6C).

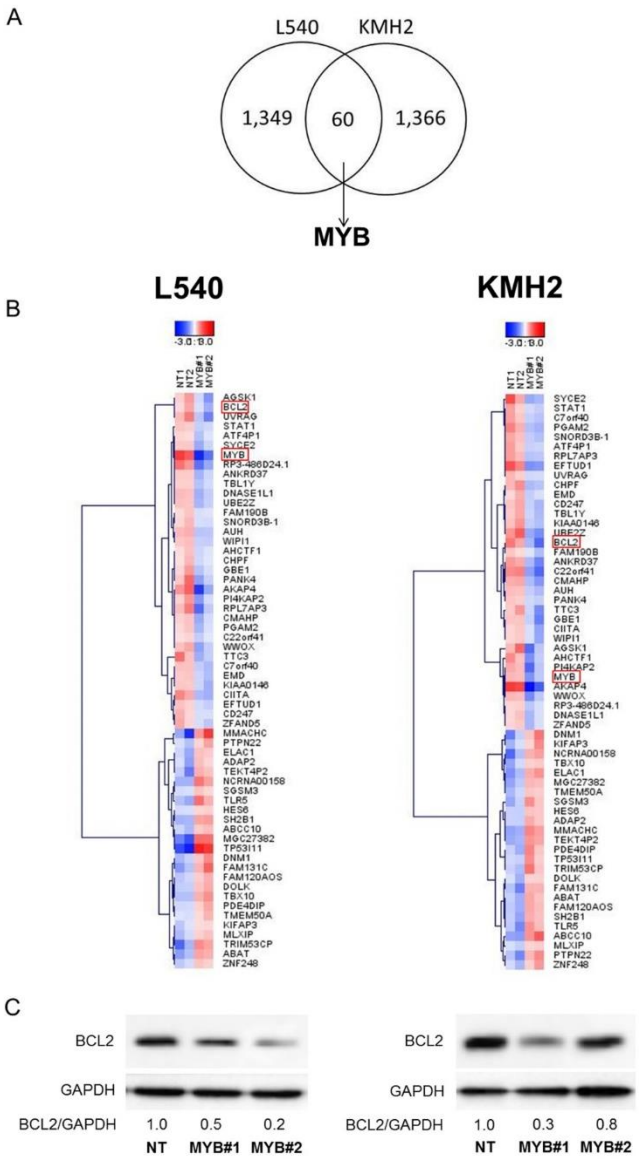



**Figure 6. Target genes of MYB in HL cell lines. (A)** Venn Diagram of the differentially expressed genes in both L540 and KMH2 cells. 60 genes are overlapped between both cell lines. 1,349 out of 1,409 in L540 and 1,366 out of 1,426 in KMH2 are genes expressed in specific cell lines. **(B)** Heatmap of the 60 genes that were overlapped between L540 and KMH2 cells. 35 of the genes were up-regulated and 25 were down-regulated. **(C)** Validation of BCL2 protein expression by western blot.

**Table 1.** 60 significantly differentially regulated genes overlap between L540 and KMH2

| Probe ID | Gene Symbol | L540 (shRNA vs NT) | | KMH2 (shRNA vs NT) | |
|---|---|---|---|---|---|
| | | Regulation | Fold | Regulation | Fold |
| A_23_P31073 | MYB | down | -3.58 | down | -1.47 |
| A_33_P3417920 | AKAP4 | down | -2.42 | down | -4.87 |
| A_33_P3369914 | RP3-486D24.1 | down | -1.95 | down | -1.24 |
| A_33_P3374903 | RPL7AP3 | down | -1.76 | down | -1.42 |
| A_33_P3260722 | UVRAG | down | -1.68 | down | -1.25 |
| A_23_P149690 | PANK4 | down | -1.59 | down | -1.26 |
| A_32_P209960 | CIITA | down | -1.59 | down | -1.32 |
| A_33_P3276913 | TTC3 | down | -1.56 | down | -1.51 |
| A_33_P3222659 | AGSK1 | down | -1.48 | down | -1.83 |
| A_32_P507710 | PI4KAP2 | down | -1.45 | down | -1.65 |
| A_24_P116606 | WWOX | down | -1.45 | down | -1.50 |
| A_32_P161166 | DNASE1L1 | down | -1.38 | down | -1.28 |
| A_23_P20852 | AUH | down | -1.38 | down | -1.35 |
| A_24_P506977 | C7orf40 | down | -1.34 | down | -1.37 |
| A_33_P3390521 | UBE2Z | down | -1.34 | down | -1.93 |
| A_23_P121082 | GBE1 | down | -1.34 | down | -1.49 |
| A_32_P203528 | SYCE2 | down | -1.34 | down | -1.81 |
| A_33_P3355185 | BCL2 | down | -1.33 | down | -1.89 |
| A_33_P3279475 | KIAA0146 | down | -1.33 | down | -1.42 |
| A_33_P3281567 | CMAHP | down | -1.32 | down | -1.56 |
| A_24_P237586 | ANKRD37 | down | -1.31 | down | -1.66 |
| A_23_P85171 | EMD | down | -1.30 | down | -1.26 |
| A_24_P290354 | ZFAND5 | down | -1.30 | down | -1.28 |
| A_23_P141394 | WIPI1 | down | -1.29 | down | -1.30 |
| A_33_P3268313 | PGAM2 | down | -1.28 | down | -1.54 |
| A_24_P754817 | EFTUD1 | down | -1.26 | down | -2.07 |
| A_23_P114466 | TBL1Y | down | -1.25 | down | -1.36 |
| A_24_P274270 | STAT1 | down | -1.25 | down | -1.30 |
| A_33_P3286873 | CD247 | down | -1.25 | down | -1.39 |
| A_33_P3226542 | SNORD3B-1 | down | -1.24 | down | -1.38 |
| A_33_P3342430 | ATF4P1 | down | -1.23 | down | -1.34 |
| A_33_P3666346 | AHCTF1 | down | -1.21 | down | -1.56 |
| A_33_P3262012 | CHPF | down | -1.21 | down | -1.48 |
| A_33_P3214314 | FAM190B | down | -1.19 | down | -1.24 |
| A_33_P3250887 | C22orf41 | down | -1.19 | down | -2.02 |
| A_23_P150281 | TP53I11 | up | 5.14 | up | 1.35 |
| A_24_P83586 | MMACHC | up | 2.79 | up | 1.67 |
| A_33_P3492750 | MGC27382 | up | 2.47 | up | 1.42 |
| A_33_P3324409 | TRIM53CP | up | 1.85 | up | 1.48 |
| A_24_P169073 | FAM131C | up | 1.71 | up | 1.56 |
| A_23_P132057 | NCRNA00158 | up | 1.66 | up | 1.74 |
| A_33_P3263423 | TLR5 | up | 1.63 | up | 1.59 |
| A_33_P3406828 | TEKT4P2 | up | 1.55 | up | 1.57 |
| A_24_P185029 | SH2B1 | up | 1.54 | up | 1.43 |
| A_33_P3268487 | ABAT | up | 1.53 | up | 1.40 |
| A_32_P538017 | TBX10 | up | 1.49 | up | 1.33 |
| A_24_P818268 | DNM1 | up | 1.43 | up | 1.67 |
| A_23_P49816 | ADAP2 | up | 1.35 | up | 1.32 |
| A_23_P15944 | ELAC1 | up | 1.35 | up | 1.63 |
| A_23_P10870 | DOLK | up | 1.34 | up | 1.28 |
| A_33_P3241984 | PTPN22 | up | 1.33 | up | 1.64 |
| A_33_P3372566 | ABCC10 | up | 1.32 | up | 1.91 |
| A_33_P3268005 | FAM120AOS | up | 1.31 | up | 1.44 |
| A_23_P9823 | MLXIP | up | 1.30 | up | 1.22 |
| A_33_P3326992 | PDE4DIP | up | 1.27 | up | 1.47 |
| A_23_P35343 | ZNF248 | up | 1.26 | up | 1.36 |
| A_24_P113824 | TMEM50A | up | 1.22 | up | 1.31 |
| A_23_P166491 | SGSM3 | up | 1.21 | up | 1.46 |
| A_23_P62920 | KIFAP3 | up | 1.20 | up | 1.41 |
| A_33_P3322864 | HES6 | up | 1.20 | up | 1.28 |

## DISCUSSION

In this study we confirmed presence of nonsense mutations in *MYB* in 2 HL cell lines (L428 and SUPHD1), which resulted in truncated MYB proteins resembling alternative splice variants of MYB. Alternative splicing was not observed in any of the HL cell lines, including the five *MYB* wild type cell lines. We previously showed that MYB is overexpressed in HL cell lines compared to normal germinal center B cells irrespective of having wild type or mutant MYB.[2] Overexpression of MYB has been shown in several types of cancers.[40-42] Moreover, overexpression of MYB has been shown to contribute to the pathogenesis of acute myeloid leukemia[43] and ovarian cancer[44]. In this study, we showed that inhibition of MYB using shRNA constructs induced negative effects on cell growth in two wild type MYB HL cell lines (L540 & KMH2) and in one truncated MYB cell line (SUPHD1). In some of the cell lines this can be explained by an increased number of apoptotic cells in MYB shRNA samples compared to the control. Thus, our findings indicate an oncogenic role of both full-length and truncated MYB protein in HL. We did not observe an effect on growth in one of the two MYB mutant cell lines. In part, this might have been caused by the lower effectiveness of the shRNA constructs as shown on the mRNA and protein levels. We do not know why the effectiveness of the shRNA constructs differs between the cell lines, as endogenous MYB mRNA and protein levels are quite similar.



Microarray analysis revealed a limited overlap in the set of MYB regulated genes in L540 and KMH2 cells. The reason for this might be the difference in cellular origin and histological subtype. L540 is most likely derived from T cells, while KMH2 originates from B cells.[45] Moreover, L540 has been derived from a nodular sclerosis subtype and KMH2 from a mixed cellularity cHL subtype. The set of target genes regulated by MYB might be cell type specific,[9] and this might explain the limited overlap. BCL2, a known downstream MYB target, suppresses apoptosis via its interaction with other pro-apoptotic members of the BCL2 family. We showed that upon MYB knockdown, BCL2 levels go down. Probably, the anti-apoptotic properties of MYB in HL are mediated at least in part by BCL2. We indeed observed an increase in the percentage of apoptotic cells upon MYB knockdown in two of the four HL cell lines.

Part of the genes that are downregulated upon MYB knockdown are not related to cell growth, e.g. CIITA and CD247. CIITA encodes for the HLA class II transactivator. Previous studies have shown that chromosomal rearrangements of CIITA are involved in the pathogenesis of HL.[46] Loss of HLA expression as a possible immune escape mechanism has been proposed for HL. CD247

encodes for the T-cell receptor zeta chain which is important for coupling antigen recognition to several intracellular signal-transduction pathways. Thus it might be that MYB expression also contributes to the cross talk with the microenvironment and possibly contributes to escape from effective anti-tumor responses. This needs to be studied in more detail to support this proposed link.

Several studies showed that MYC might be a potential target gene of MYB based on presence of MYB binding sites in the promotor region of MYC and on MYB reporter assays.[35,36] A direct regulation was also confirmed by chromatin immunoprecipitation (ChIP) assays.[47,48] In HL, we did not see a reduction of MYC protein levels upon MYB knockdown. This might be caused by the complex regulation of the *MYC* gene by many regulatory elements.[49,50] Other studies have suggested an opposite regulation, in which MYC regulates expression of MYB, possibly by repressing the expression of miR-150.[17] In addition, He et al.[37] showed that MYC is able to inhibit the maturation of miR-150 and thereby indirectly regulated expression of MYB. The levels of miR-150 are very low or even undetectable in wild type and MYB/MYC shRNA treated HL cell lines, making a miR-150 mediated regulation of MYB in HL unlikely. Knockdown of MYC in HL cell lines did result in reduced levels of MYB protein. Thus, based on the data that we have, MYB seems to be regulated by MYC independently of miR-150. MYC knockdown inhibited proliferation of HL cells, independently of presence of a mutation in the *MYB* gene. The effect of MYC knockdown was more pronounced as compared to the effect of MYB knockdown, which might indicate that MYC has a more dominant role on growth of HL cells compared to MYB.

In conclusion, we found strong oncogenic effects of wild type and truncated MYB in HL cell lines. Gene expression profiling of HL cells upon MYB knockdown revealed a high number of MYB regulated genes, although the overlap between the two cell lines was quite limited. MYB shRNA inhibited HL cell proliferation through apoptosis induction probably via BCL2. MYB protein levels are regulated by MYC, whereas MYC levels are independent of MYB in HL.

**ACKNOWLEGMENTS**

This work was supported by grants from the Abel Tasman Talent Program of the University Medical Center Groningen.

**Supplementary Information**

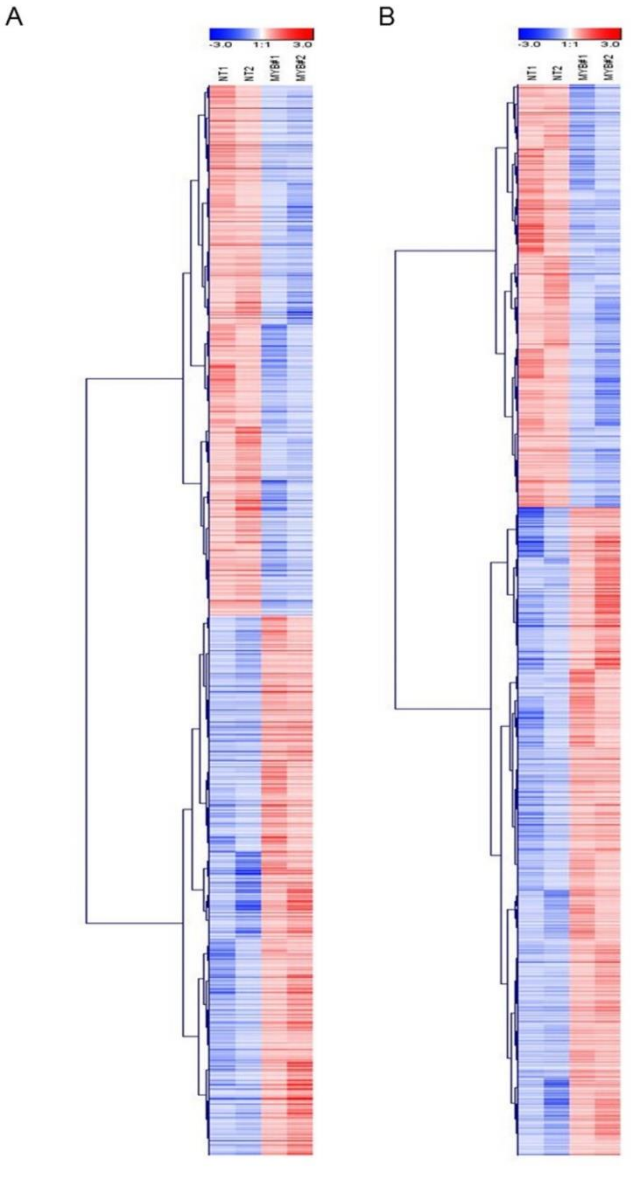



**Figure S1. Heatmaps of the unsupervised hierarchical clustering of the probes with a significantly different expression level using a moderate T-test. (A)** Heatmap of L540 cells including 1,474 differentially expressed probes (1,409 genes) with 50% of the genes being up-regulated **(B)** Heatmap of KMH2 cells, 61% of the 1,481 differentially expressed probes (1,426 genes) being up-regulated.

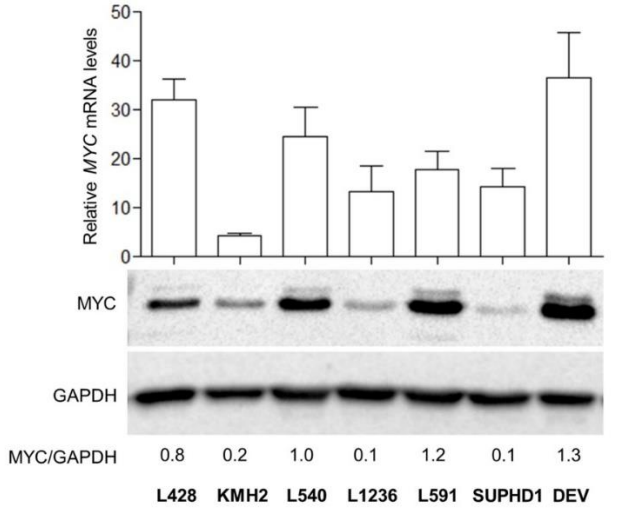



**Figure S2. Endogenous MYC levels in HL cell lines. (A)** *MYC* mRNA levels by qRT-PCR (mean with SD) and **(B)** MYC protein by western blot in HL cell lines. The effect we observed in the GFP competition assays is not dependent on the endogenous MYC levels in HL cell lines.

**Table S1.** Primer sets and oligos

| Primer & Oligo | Forward (5' to 3') | Reverse (5' to 3') | Amplicon (bp) |
|---|---|---|---|
| primer sequences (alternative exon) | | | |
| MYB-L428/RNA | GGTGGCACAGCACCACCATT | ATGGTGCCCTGGTGGACGAT | 169 |
| MYB-SUPHD1/DNA | GCCTTCTTTAACTTCCACCC | GTTGTAATATGCCCAGGACAG | 270 |
| MYB-SUPHD1/RNA | GCCTTCTTTAACTTCCACCC | CTTGAGCTGCAAGTGCATGT | 185 |
| MYB-8A | AAGGACAGCAGGTGCTACCA | AACAGGTGCACTGTCTCCAT | 107 |
| MYB-9A/9B | CCAGCAAGGTGCATGATCGT | TTTCACAGTCTGGTCTCTA | 216 |
| MYB-10A | AAATGCCTTCTTTAACTTCC | CTTGAGCTGCAAGTGCATGT | 189 |
| MYB-13A | CCAACTGTTCACGCAGACCT | CTTCTGATGCTGGTGCCATT | 83 |
| MYB-14A | ATGGCACCAGCATCAGAAGA | CGAGCTTGACTGGAAGATGT | 158 |
| RPL22 | TCGCTCACCTCCCTTTCTAA | TCACGGTGATCTTGCTCTTG | 250 |
| primer sequences (qPCR) | | | |
| MYB | CCAACTGTTCACGCAGACCT | CTTCTGATGCTGGTGCCATT | 83 |
| MYC | CACCAGCAGCGACTCTGA | ATCCAGACTCTGACCTTTTGC | 101 |
| TBP | GCCCGAAACGCCGAATAT | CCGTGGTTCGTGGCTCTCT | 73 |
| shRNA sequences | | | |
| MYB#1 S | GATCCGGTTATCTGCAGGAGTCTTCTTCAAGAGAGAAGACTCCTGCAGATAACCTTTTTG | | |
| MYB#1 AS | AATTCAAAAAGGTTATCTGCAGGAGTCTTCTCTCTTGAAGAAGACTCCTGCAGATAACCG | | |
| MYB#2 S | GATCCGACACCCTCTCATCTAGTAGTTCAAGAGACTACTAGATGAGAGGGTGTCTTTTTG | | |
| MYB#2 AS | AATTCAAAAAGACACCCTCTCATCTAGTAGTCTCTTGAACTACTAGATGAGAGGGTGTCG | | |
| MYC#1 S | GATCCGATGAGGAAGAAATCGATGTTCAAGAGACATCGATTTCTTCCTCATCTTTTTG | | |
| MYC#1 AS | AATTCAAAAAGATGAGGAAGAAATCGATGTCTCTTGAACATCGATTTCTTCCTCATCG | | |
| MYC#2 S | GATCCAACGACGAGAACAGTTGAAACATTCAAGAGATGTTTCAACTGTTCTCGTCGTTTTTTTG | | |
| MYC#2 AS | AATTCAAAAAAACGACGAGAACAGTTGAAACATCTCTTGAATGTTTCAACTGTTCTCGTCGTTG | | |

S, sense; AS, antisense

## REFERENCES


1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013; 49(6): 1374-403.
2. Liu Y, Abdul Razak FR, Terpstra M, Chan FC, Saber A, Nijland M et al. The mutational landscape of Hodgkin lymphoma cell lines determined by whole-exome sequencing. Leukemia 2014; 28(11): 2248-51.
3. Zhou Y, Ness SA. Myb proteins: angels and demons in normal and transformed cells. Front Biosci (Landmark Ed) 2011; 16: 1109-31.
4. Klempnauer KH, Gonda TJ, Bishop JM. Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: the architecture of a transduced oncogene. Cell 1982; 31(2 Pt 1): 453-63.
5. Leprince D, Gegonne A, Coll J, de Taisne C, Schneeberger A, Lagrou C et al. A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. Nature 1983; 306(5941): 395-7.
6. Gonda TJ, Bishop JM. Structure and transcription of the cellular homolog (c-myb) of the avian myeloblastosis virus transforming gene (v-myb). J Virol 1983; 46(1): 212-20.
7. Gonda TJ, Buckmaster C, Ramsay RG. Activation of c-myb by carboxy-terminal truncation: relationship to transformation of murine haemopoietic cells in vitro. Embo J 1989; 8(6): 1777-83.
8. Dash AB, Orrico FC, Ness SA. The EVES motif mediates both intermolecular and intramolecular regulation of c-Myb. Genes Dev 1996; 10(15): 1858-69.
9. Lei W, Rushton JJ, Davis LM, Liu F, Ness SA. Positive and negative determinants of target gene specificity in myb transcription factors. J Biol Chem 2004; 279(28): 29519-27.
10. Favier D, Gonda TJ. Detection of proteins that bind to the leucine zipper motif of c-Myb. Oncogene 1994; 9(1): 305-11.
11. Nomura T, Tanikawa J, Akimaru H, Kanei-Ishii C, Ichikawa-Iwata E, Khan MM et al. Oncogenic activation of c-Myb correlates with a loss of negative regulation by TIF1beta and Ski. J Biol Chem 2004; 279(16): 16715-26.
12. Tavner FJ, Simpson R, Tashiro S, Favier D, Jenkins NA, Gilbert DJ et al. Molecular cloning reveals that the p160 Myb-binding protein is a novel, predominantly nucleolar protein which may play a role in transactivation by Myb. Mol Cell Biol 1998; 18(2): 989-1002.
13. Bies J, Wolff L. Oncogenic activation of c-Myb by carboxyl-terminal truncation leads to decreased proteolysis by the ubiquitin-26S proteasome pathway. Oncogene 1997; 14(2): 203-12.
14. George OL, Ness SA. Situational awareness: regulation of the myb transcription factor in differentiation, the cell cycle and oncogenesis. Cancers (Basel) 2014; 6(4): 2049-71.
15. Lahortiga I, De Keersmaecker K, Van Vlierberghe P, Graux C, Cauwelier B, Lambert F et al. Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia. Nat Genet 2007; 39(5): 593-5.
16. Clappier E, Cuccuini W, Kalota A, Crinquette A, Cayuela JM, Dik WA et al. The C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia (T-ALL), the translocation defining a new T-ALL subtype in very young children. Blood 2007; 110(4): 1251-61.
17. Persson M, Andren Y, Mark J, Horlings HM, Persson F, Stenman G. Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. Proc Natl Acad Sci U S A 2009; 106(44): 18740-4.

**18.** O'Rourke JP, Ness SA. Alternative RNA splicing produces multiple forms of c-Myb with unique transcriptional activities. Mol Cell Biol 2008; 28(6): 2091-101.
**19.** Shen-Ong GL, Skurla RM, Jr., Owens JD, Mushinski JF. Alternative splicing of RNAs transcribed from the human c-myb gene. Mol Cell Biol 1990; 10(6): 2715-22.
**20.** Westin EH, Gorse KM, Clarke MF. Alternative splicing of the human c-myb gene. Oncogene 1990; 5(8): 1117-24.
**21.** Shen-Ong GL. Alternative internal splicing in c-myb RNAs occurs commonly in normal and tumor cells. Embo J 1987; 6(13): 4035-9.
**22.** Schuur ER, Dasgupta P, Reddy EP, Rabinovich JM, Baluda MA. Alternative splicing of the chicken c-myb exon 9A. Oncogene 1993; 8(7): 1839-47.
**23.** Woo CH, Sopchak L, Lipsick JS. Overexpression of an alternatively spliced form of c-Myb results in increases in transactivation and transforms avian myelomonoblasts. J Virol 1998; 72(8): 6813-21.
**24.** Brill LB, 2nd, Kanner WA, Fehr A, Andren Y, Moskaluk CA, Loning T et al. Analysis of MYB expression and MYB-NFIB gene fusions in adenoid cystic carcinoma and other salivary neoplasms. Mod Pathol 2011; 24(9): 1169-76.
**25.** Frykberg L, Metz T, Brady G, Introna M, Beug H, Vennstrom B et al. A point mutation in the DNA binding domain of the v-myb oncogene of E26 virus confers temperature sensitivity for transformation of myelomonocytic cells. Oncogene Res 1988; 3(4): 313-22.
**26.** Introna M, Golay J, Frampton J, Nakano T, Ness SA, Graf T. Mutations in v-myb alter the differentiation of myelomonocytic cells transformed by the oncogene. Cell 1990; 63(6): 1289-97.
**27.** Stober-Grasser U, Brydolf B, Bin X, Grasser F, Firtel RA, Lipsick JS. The Myb DNA-binding domain is highly conserved in Dictyostelium discoideum. Oncogene 1992; 7(3): 589-96.
**28.** Dini PW, Eltman JT, Lipsick JS. Mutations in the DNA-binding and transcriptional activation domains of v-Myb cooperate in transformation. J Virol 1995; 69(4): 2515-24.
**29.** Fu SL, Lipsick JS. FAETL motif required for leukemic transformation by v-Myb. J Virol 1996; 70(8): 5600-10.
**30.** Leverson JD, Ness SA. Point mutations in v-Myb disrupt a cyclophilin-catalyzed negative regulatory mechanism. Mol Cell 1998; 1(2): 203-11.
**31.** Ramsay RG, Gonda TJ. MYB function in normal and cancer cells. Nat Rev Cancer 2008; 8(7): 523-34.
**32.** Zorbas M, Sicurella C, Bertoncello I, Venter D, Ellis S, Mucenski ML et al. c-Myb is critical for murine colon development. Oncogene 1999; 18(42): 5821-30.
**33.** Thompson MA, Rosenthal MA, Ellis SL, Friend AJ, Zorbas MI, Whitehead RH et al. c-Myb down-regulation is associated with human colon cell differentiation, apoptosis, and decreased Bcl-2 expression. Cancer Res 1998; 58(22): 5168-75.
**34.** Alitalo K, Winqvist R, Lin CC, de la Chapelle A, Schwab M, Bishop JM. Aberrant expression of an amplified c-myb oncogene in two cell lines from a colon carcinoma. Proc Natl Acad Sci U S A 1984; 81(14): 4534-8.
**35.** Nakagoshi H, Kanei-Ishii C, Sawazaki T, Mizuguchi G, Ishii S. Transcriptional activation of the c-myc gene by the c-myb and B-myb gene products. Oncogene 1992; 7(6): 1233-40.
**36.** Cogswell JP, Cogswell PC, Kuehl WM, Cuddihy AM, Bender TM, Engelke U et al. Mechanism of c-myc regulation by c-Myb in different cell lineages. Mol Cell Biol 1993; 13(5): 2858-69.
**37.** He Y, Jiang X, Chen J. The role of miR-150 in normal and malignant hematopoiesis. Oncogene 2014; 33(30): 3887-93.

**38.** Atayar C, Kok K, Kluiver J, Bosga A, van den Berg E, van der Vlies P et al. BCL6 alternative breakpoint region break and homozygous deletion of 17q24 in the nodular lymphocyte predominance type of Hodgkin's lymphoma-derived cell line DEV. Hum Pathol 2006; 37(6): 675-83.
**39.** Li J, Kluiver J, Osinga J, Westers H, van Werkhoven MB, Seelen MA et al. Functional Studies on Primary Tubular Epithelial Cells Indicate a Tumor Suppressor Role of SETD2 in Clear Cell Renal Cell Carcinoma. Neoplasia 2016; 18(6): 339-46.
**40.** Srivastava SK, Bhardwaj A, Singh S, Arora S, McClellan S, Grizzle WE et al. Myb overexpression overrides androgen depletion-induced cell cycle arrest and apoptosis in prostate cancer cells, and confers aggressive malignant traits: potential role in castration resistance. Carcinogenesis 2012; 33(6): 1149-57.
**41.** Thorner AR, Parker JS, Hoadley KA, Perou CM. Potential tumor suppressor role for the c-Myb oncogene in luminal breast cancer. PLoS One 2010; 5(10): e13073.
**42.** Biroccio A, Benassi B, D'Agnano I, D'Angelo C, Buglioni S, Mottolese M et al. c-Myb and Bcl-x overexpression predicts poor prognosis in colorectal cancer: clinical and experimental findings. Am J Pathol 2001; 158(4): 1289-99.
**43.** Uttarkar S, Dasse E, Coulibaly A, Steinmann S, Jakobs A, Schomburg C et al. Targeting acute myeloid leukemia with a small molecule inhibitor of the Myb/p300 interaction. Blood 2016; 127(9): 1173-82.
**44.** Barletta C, Lazzaro D, Prosperi Porta R, Testa U, Grignani F, Ragusa RM et al. C-MYB activation and the pathogenesis of ovarian cancer. Eur J Gynaecol Oncol 1992; 13(1): 53-9.
**45.** Hanamoto H, Nakayama T, Miyazato H, Takegawa S, Hieshima K, Tatsumi Y et al. Expression of CCL28 by Reed-Sternberg cells defines a major subtype of classical Hodgkin's disease with frequent infiltration of eosinophils and/or plasma cells. Am J Pathol 2004; 164(3): 997-1006.
**46.** Steidl C, Shah SP, Woolcock BW, Rui L, Kawahara M, Farinha P et al. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. Nature 2011; 471(7338): 377-81.
**47.** Berge T, Matre V, Brendeford EM, Saether T, Luscher B, Gabrielsen OS. Revisiting a selection of target genes for the hematopoietic transcription factor c-Myb using chromatin immunoprecipitation and c-Myb knockdown. Blood Cells Mol Dis 2007; 39(3): 278-86.
**48.** Ciznadija D, Tothill R, Waterman ML, Zhao L, Huynh D, Yu RM et al. Intestinal adenoma formation and MYC activation are regulated by cooperation between MYB and Wnt signaling. Cell Death Differ 2009; 16(11): 1530-8.
**49.** Shi J, Whyte WA, Zepeda-Mendoza CJ, Milazzo JP, Shen C, Roe JS et al. Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. Genes Dev 2013; 27(24): 2648-62.
**50.** Wright JB, Brown SJ, Cole MD. Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells. Mol Cell Biol 2010; 30(6): 1411-20.

# Chapter 6

## Summary, Discussion and Future perspective

## SUMMARY AND DISCUSSION

The tumor cells in Hodgkin lymphoma (HL) are characterized by constitutive activation of several signaling pathways and expression of a distinct set of transcription factors. This characteristic phenotype is partly caused by gene mutations as determined largely by targeted sequencing approaches on cell lines and microdissected Hodgkin and Reed-Sternberg (HRS) cells. The aim of this thesis was to identify additional mutations by a more comprehensive sequencing approach and to substantiate the functional importance of some of these mutations in HL.

### *Mutational landscape of HL*

In recent years, high-throughput sequencing technology has provided novel opportunities for the comprehensive identification of genetic aberrations involved in various types of cancer. The first objective of this thesis was to identify commonly mutated genes in HL. In **chapter 2,** we determined the mutational landscape of HL cell lines by a whole exome sequencing (WES) approach. Overall, we identified 463 recurrently mutated genes in HL regardless of subtype and 373 recurrent mutations specifically in classical (c)HL. The two most common gene ontologies within the mutated genes are cell adhesion and cell development and differentiation. Among 373 genes mutated in cHL, 44 showed a significantly different expression pattern compared to germinal center B cells. Several of the commonly mutated genes have also been found to be mutated in other germinal center B cell derived lymphoma subtypes. Part of the consistently mutated genes map to regions with copy number gain or loss. Overall, our data together with findings of other research groups support the potential relevance of these genes in HL pathogenesis and help in the selection of candidate genes for further functional study.



Due to the scarcity of the malignant HRS cells in diagnostic biopsies it is challenging to isolate purified HRS cells in sufficiently high numbers for genome-wide analyses. Recently, Reichel et al.[1] successfully performed genome-wide analyses on primary HRS cells using fluorescence activated cell sorting (FACS) to isolate HRS cells from their microenvironment. They used flow-sorted HRS cells from 10 HL patients and identified 99 recurrently mutated genes. Comparison of our data to their findings revealed 14 commonly mutated genes in both studies. These genes are ABCA13, B2M, CSF2RB, DMD, DNAH14, FAT4, FSIP4 GPR98, ITPKB, MLL3, PCLO, TNFAIP3, TRMT2A and ZNF217.

Several genes previously reported to be commonly mutated, e.g. members of the NF-KB and JAK/STAT pathways, were also found to be mutated in the flow-sorted HRS cells and HL cell lines.[1,2] The most commonly altered genes are TNFAIP3 in 6 of 10 cases and SOCS1 in 4 of 10 cases.[1] TNFAIP3 is involved in the negative regulation of the NF-κB signalling pathway. TNFAIP3 is frequently inactivated by gene deletions/mutations in a variety of B-cell malignancies.[3] Multiple members of this pathway, such as REL, NIK, MAP3K14, NFKBIA, NFKB1E, TRAF3 and CYLD, were found to be commonly mutated in HL in previous studies. Overall, these findings show that mutations found in HL cell lines were also present in primary tumors. Results from these genome-wide analyses will help to find novel genes that are potentially relevant for HL pathogenesis.

***Mutations in immune system related genes in HL***

HRS cells are rare in the involved lymph nodes and they reside within an abundant inflammatory microenvironment, largely consisting of various T cell subsets. Chemokines and cytokines produced by HRS cells and by the infiltrating cells shape the environment and provide proliferative and survival signals to the HRS cells. Despite the critical dependence on the microenvironment, HRS cells need to apply mechanisms to evade cytotoxic T cell (CTL) and natural killer (NK) cell mediated anti-tumor responses. These mechanisms include secretion of immune-suppressive factors like IL10 and TGFβ, recruitment of regulatory and helper T cells, expression of PDL1 and CD95 and loss of HLA expression. Recent publications show that mutations in immune system related genes may represent a mechanism of HRS cells to evade detection by immune cells.[4] Several reports have shown lack of expression of HLA classes I and II by HRS cells.[5-7] Expression of HLA classes I and II is important for recognition of antigenic peptides by the T cells. In **chapter 2**, we showed *B2M* mutations affecting the ATG start codon in L428 (heterozygous) and DEV (homozygous) cells. Consistent with that, we observed no or very low membrane B2M and HLA class I expression in L428 and DEV by flow cytometry. *B2M* mRNA levels were reduced in both cell lines as compared to L1236, whereas HLA-A, HLA-B and HLA-C levels were in the same range. WES of flow sorted HRS cells showed that B2M is the most commonly altered gene in HRS cells.[1] From these findings it appears that inactivating mutations in B2M are the major cause of loss of MHC class I expression in HL, which provides an immune escape mechanism to HRS cells.

In addition, we observed *CIITA* mutations in 2 of the 7 HL cell lines albeit with low read counts. These mutations were confirmed by Sanger sequencing and were also found in the Broad data set. CIITA is a master regulator of the HLA class II promoter activity which often cause unbalanced expression of HLA-II.[8,9] It was already shown that CIITA gene is inactivated by a chromosomal translocation in cHL.[8] Interestingly, we also found mutations in CREB binding protein (CREBBP) in 2 HL cell lines (L540 and SUPHD1). CREBBP promotes CIITA-dependent transcription including HLA class II.[10] Recently, CREBBP mutations have been associated with decreased antigen presentation and reduced MHC class II expression on tumor cells in follicular lymphoma.[11] In flow sorted HRS cells Reichel et al.[1] also found mutations in CIITA and CREBBP in 1 out of 10 cases. From our findings and together with previously reported data, inactivation of CIITA and mutations of CREBBP might very well explain downregulation of the HLA class II expression which is frequently observed in primary cHL cases.

In **chapter 3**, we focused on mutations in CD58 and studied its expression in HL cell lines and primary HL tissue samples. CD58 (LFA-3) is an adhesion molecule that is involved in immune recognition of tumor cells by binding to the CD2 receptor expressed on cytotoxic T cells. WES analysis revealed a stop gain mutation of *CD58* in DEV, a splice donor site mutation in KMH2 and loss of exon 1-3 in SUPHD1. *CD58* mutations were confirmed in all three HL cell lines by Sanger sequencing at the DNA and RNA level. *CD58* mRNA levels were low or absent in SUPHD1 and KMH2 cells and normal in DEV. CD58 protein expression as determined by flow, western blot and IHC was absent in all 3 mutated HL cell lines. In four HL cell lines with wild type CD58 we did observe CD58 protein expression. These data confirm presence of *CD58* gene mutations and indicate that these mutations lead to loss of CD58 protein expression in 3 out of 7 HL cell lines. Another group also observed CD58 mutations in SUPHD1, KMH2 and in a third HL cell line, UHO-1, which we did not include in our analysis.[12] They identified heterozygous deletions of CD58 by FICTION analysis in 3 out of 13 primary cHL cases. Sequencing analysis of the CD58 gene on DNA isolated from microdissected HRS cells of 10 cHL cases did not reveal any mutations. Reichel et al.[1] showed CD58 gene deletions in flow-isolated HRS cells in 2 out of 10 cHL cases. Based on these studies and on our data, it is obvious that alterations of the CD58 gene are common in HL cell lines and are also present in primary HRS cells of patients.

We next set out to define whether loss of CD58 protein expression is common in HRS cells of primary HL tumor samples. In contrast to the cell lines, tumor cells of 43 primary HL cases with good treatment outcome all showed strong

CD58 expression. As HL cell lines are derived from end stage HL patients, we next studied CD58 expression in primary and relapsed tissue samples of relapsed HL patients. Loss of CD58 staining was observed in HRS cells in 6 out of 53 patients who experienced a relapse. These findings indicate that loss of CD58 on tumor cells might be restricted to HL patients who do experience relapse. Loss of CD58 expression may possibly occur as a late event during progression and might happen when HRS cells evolve to become independent of the T-cell infiltrate for survival. At that point, loss of CD58 might be a prerequisite for survival of tumor cells as the tumor cells have accumulated mutations that can potentially lead to presentation of highly immunogenic peptides that can activate cytotoxic T cell responses. Our results indicate that mutations in *CD58* and loss of CD58 expression are common in HL derived cell lines and that loss of CD58 expression in tumor cells is possibly a late event associated with relapse of disease. These data suggest that loss of CD58 is a potential immune escape mechanism of HL tumor cells, especially in clinically aggressive disease. In diffuse large B cell lymphoma (DLBCL), mutations of the *CD58* gene are reported to contribute to immune evasion of the tumor cells. Challa-Malladi et al.[13] reported that loss of CD58 expression is associated with loss of HLA class I in DLBCL and this co-loss is crucial to avoid triggering of NK cell activation by CD58. In our studies, loss of CD58 expression is not restricted to HLA class I negative HL cell lines. Due to the low number of CD58 negative cases in tissue, we cannot conclude whether loss of CD58 is specific for HLA class I negative cases.

In addition to mutations in *CD58*, we also found mutations in *CSF2RB*, another immune associated gene, in HL cell lines. The *CSF2RB* gene, encoding the common β chain (CD131) shared by the interleukin-3 (IL-3), granulocytic macrophage colony-stimulating factor (GM-CSF) and IL-5 receptors, was mutated in 4 out of 7 HL cell lines.[2] In **chapter 4**, our preliminary data confirmed the two missense mutations (V212I and A715S) in KMH2, a missense (V524M) and frameshift mutation (c.2500_2501del) in SUPHD1, a frameshift mutation (c.2198del) in L1236 and a stop gain mutation (Q809*) in DEV cells by Sanger-sequencing at DNA and mRNA levels. Despite several attempts we were unable to amplify the region covering both mutations in KMH2 and SUPHD1 at the DNA or RNA level. Thus we were unable to determine whether the mutations were present at the same parental allele or not. Reichel et al.[1] also found CSF2RB mutations in 3 out of 10 cHL cases. This indicates that mutations not only occur in HL cell lines, but also in primary cases. In the HL cell lines we failed to induce growth by addition of exogenous cytokines. As JAK2/STAT5 is the main signalling pathway for CSF2RB to promote survival, proliferation and differentiation, the reason might be related

to somatic mutations or amplification of SOCS1, PTPN2 or JAK2. We did not yet look at the expression of JAK2, STAT5 and pSTAT5, which might help to design further studies to unravel the role of CSF2RB and the consequences of the mutations in HL.

***Mutations in the transcription factor MYB***

The *MYB* proto-oncogene encodes for the transcriptional activator MYB, which is involved in the regulation of proliferation and differentiation of hematopoietic cells. In our WES data, we found frame shift mutations in *MYB* in 2 out of 7 HL cell lines. In **chapter 5**, we confirmed the *MYB* mutations in 2 HL cell lines at the DNA and RNA level by Sanger sequencing. In addition to mutations, alternative splicing can also result in truncated MYB protein variants. Analysis of alternative splicing in all 7 HL cell lines with primers flanking each known alternatively spliced exon indicated that there were no alternatively spliced transcripts detectable in the HL cell lines. The frame shift mutation in L428 causes loss of most of the C-terminal domain and the mutation in SUPHD1 results in a partial loss of the C-terminal domain. We confirmed presence of the truncated protein by Western blot with an N-terminal domain specific antibody and confirmed loss of full-length MYB protein by Western blot with a C-terminus specific antibody.



Staining of MYB in HL cases revealed no staining with both antibodies in the vast majority of the cases, indicating complete loss of the MYB protein. As HL cell lines are derived from refractory or relapsed HL disease, it might be that overexpression of MYB is a late event in HL progression and may occur when HRS cells become independent of the reactive infiltrate and need stronger endogenous proliferation signals for cell survival.

In chapter 2 we showed that *MYB* is overexpressed in HL cell lines compared to normal germinal center B cells.[2] Overexpression of MYB has also been shown in several other types of cancers.[14-16] Moreover, overexpression of MYB has been show to contribute to the pathogenesis of acute myeloid leukemia (AML)[17] and ovarian cancer[18]. In these studies, we showed that inhibition of MYB using shRNA constructs induced negative effects on cell growth in two wild type MYB HL cell lines (L540 & KMH2) and in one truncated MYB cell line (SUPHD1) with an increased number of apoptotic cells in MYB shRNA samples compared to the control. Thus, our findings indicate an oncogenic role of full-length and truncated MYB protein in HL. No effect was observed in the second HL cell line with a truncating MYB mutation (L428). This might be consistent with the lower effectiveness of the shRNAs as shown on the mRNA and protein

levels. However, we do not know why the effectiveness of the shRNA constructs differs between the cell lines, as endogenous MYB levels are quite similar.

Microarray analyses of L540 & KMH2 cells treated with non-targeting and *MYB* shRNA constructs revealed 25 consistently MYB induced and 35 consistently MYB repressed genes. Gene ontology analysis indicated that the gene functions are related to proliferation, apoptosis, cell adhesion and immune response. Among these genes, we found BCL2, STAT1 and KIFAP3 as possible downstream MYB targets that may be associated with the reduced growth. BCL2, a known downstream MYB target, suppresses apoptosis via its interaction with other pro-apoptotic members of the BCL2 family. We showed that upon MYB knockdown, BCL2 levels go down. As BCL2 is a well-known anti-apoptotic protein, it can be anticipated that the anti-apoptotic properties of MYB in HL are mediated at least in part by BCL2. We indeed observed an increase in the percentage of apoptotic cells upon MYB knockdown in HL.

Several studies stated that *MYC* might be a potential target gene of MYB based on the presence of MYB binding sites in the promotor region of MYC. In HL, we did not see a reduction of MYC protein levels upon MYB knockdown. This might be caused by the complex regulation of the *MYC* gene by many regulatory elements.[19-21] On the other hand, we observed an opposite regulation, in which knockdown of MYC showed reduced MYB protein levels. Other studies have suggested that MYC regulates expression of MYB possibly by repressing the expression of microRNA-150 (miR-150).[22] The effect of MYC knockdown on proliferation of HL cells was more pronounced as compared to the effect of MYB knockdown, which indicates that MYC has a more dominant role in growth of HL cells as compared to MYB. Based on the data that we have, MYB seems to be regulated by MYC, whereas MYC levels are independent of MYB in HL. Further analyses need to be done to reveal the precise interactions between MYC and MYB in HL.

## FUTURE PERSPECTIVES

With the discovery of recurrent gene mutations and deregulated signaling pathways, a better understanding of HL pathogenesis is evolving. This knowledge might contribute to the identification and design of new therapeutic targets.

### *Exploring the role of other recurrent mutations in HL*

In this thesis, we identified 463 recurrently mutated genes in HL regardless of subtype and 373 recurrent mutations specifically in cHL. We focused functional follow-up studies on CD58, CSF2RB and MYB based on the gene ontology and their role in cancer. From the list of recurrently mutated genes, a number of other genes might be of interest to study in more detail in HL pathogenesis. Especially genes with mutations that are likely to affect the protein structure, which have also been found to be mutated in primary HL tissues are interesting, for example ABCA13, DMD, DNAH14, FAT4, FSIP4, GPR98, ITPKB, MLL3, PCLO, TNFAIP3, TRMT2A and ZNF217.[1] Several of the identified genes or gene families have not yet been linked to HL pathogenesis. From our analysis, cell adhesion was the most common gene ontology among the mutated genes and FAT atypical cadherin 4 (FAT4) is one of them. The FAT family encoding transmembrane proteins are frequently mutated in multiple human cancer types. FAT family members have growth and invasion suppressive properties in several cancer cell lines, although the mechanisms remain unknown.[23] FAT4 expression is repressed in breast cancer and lung cancer due to promoter hypermethylation.[24] Recent studies also found recurrent mutations in FAT4 in mantle cell lymphoma,[25] splenic marginal zone lymphoma[26] and diffuse large B-cell lymphoma[27]. Thus, mutations in FAT4 seem to play an important role in cancer, including B cell lymphomas.



Cell development & differentiation is also a common gene ontology among mutated genes and one of these genes is piccolo presynaptic cytomatrix protein (PCLO). PCLO encodes a protein that functions as part of the presynaptic cytoskeletal matrix thought to be involved in regulating neurotransmitter release.[28] Somatic mutations of PCLO were identified in 17 (35%) primary DLBCL cases[29] and >10% of Richter syndrome patients[30]. Zinc-finger protein 217 (ZNF217) has been proposed to act as an oncogenic protein in various human cancers such as colorectal cancer[31], gastric carcinoma[32] and ovarian cancer[33]. ZNF217 is a member of the Kruppel-like family of transcription factors and contains 8 predicted C2H2 zinc finger motifs and a proline-rich region. ZNF217 interferes with several intracellular signaling networks for reprogramming cancer cells such as proliferation, apoptosis and

invasion and metastasis. Inositol-Trisphosphate 3-Kinase B (ITPKB) regulates inositol phosphate metabolism by phosphorylation of second messenger inositol 1,4,5-trisphosphate to Ins(1,3,4,5)P4. Thus this protein is responsible for regulating the levels of a large number of inositol polyphosphates that are important in cellular signaling. A recent study reported a novel tumor suppressor role of ITPKB based on the development of DLBCL in ITPKB deficient mice.[34] ITPKB mutations were also identified in primary mediastinal B-cell lymphoma (PMBL).[35] The first step to do is to determine the protein expression of candidate genes in HRS cells by immunohistochemistry, especially for genes harbouring truncating mutations. Then, to study the functional role of these genes in HL pathogenesis, it will be important to up- or downregulate these genes and study the induced phenotype changes. GFP competition and apoptosis assays can be done to understand the oncogenic role of these genes in HL.

Previous genome wide association studies showed a strong association of the HLA region to cHL.[36] It might be interesting to analyze presence of mutations in HLA genes as a possible mechanism to escape from anti-tumor immune responses for HRS cells. As the HLA genes are highly polymorphic, we need a specific pipeline to analyze the WES data.

***RNA-seq in HL***

Although large-scale efforts for molecular profiling of cancer samples provide a broad range of data, most candidate cancer genes have been identified based on somatic mutations and DNA copy number alterations. It is, however, evident that aberrant expression of protein coding and noncoding genes might also contribute to the pathogenesis of HL. Moreover, gene fusions as the result of specific chromosomal aberrations might also contribute to disease pathogenesis. Transcriptome sequencing of HL cell lines revealed a novel recurrent gene fusion involving the CIITA and PDL1 / PDL2 genes.[8] The same group also reported 2 additional novel fusion transcripts: KIAA1432-CLDN14 (L428) and PDCD1LG2-IGHV7-81 (L1236).[37] Van Roosbroeck et al.[38] identified PDL1/2 rearrangements in four out of 200 HL cases. 25% of the HL cases has a high level amplification, including four cases with a selective amplification of PDL1/2.

The impact of noncoding (nc)RNAs in HL has been described for miRNAs.[39] Deregulated miRNA expression is reported in various human diseases including HL, suggesting an important role in their pathogenesis.[40] MicroRNAs have been shown to regulate gene expression at both the transcriptional and

translational level. Some of the predicted target genes of the upregulated miRNAs in HL are members of the SOCS (Suppressor of Cytokine signaling) family. This suggests that miRNAs may contribute to activation of the JAK/STAT signalling in HRS cells due to miRNA-mediated SOCS inactivation. Differences in miRNA expression profiles of HRS cells of primary cHL patients[41] and HL cell line[42-44] highlight the disparity between the expression profile of the tumorous HRS cell and the nonneoplastic surrounding cells. Plasma miRNA levels might be used as disease response biomarkers in cHL.[45] Thus it is important to further study the role of miRNAs in HL pathogenesis by small RNA sequencing combined with target gene identification.

Recently, a number of studies have shown that lncRNA expression can be deregulated in human cancers and they play critical roles in tumorigenesis and tumor progression in various types of cancer.[46,47] LncRNAs are a type of ncRNA that are at least 200 bp long and lack functional open reading frames. The discovery of novel lncRNAs is challenging for a number of reasons, including their frequent low expression and determination of their coding potential. However, these challenges are being overcome and several groups have performed systematic analysis of lncRNAs in normal cells and also in primary tumors.[48,49] Previous study by Verma et al.[50] reported a large number of novel lncRNAs in DLBCL. Recently, our group identified 475 differentially expressed lncRNAs, with almost 75% of them being down-regulated in cHL cell lines.[51] Moreover, we showed a dynamic regulation of lncRNA expression during the transition of B cells through the GC. These GC specific lncRNAs might be involved in the pathogenesis of HL, based on the potential high risk oncogenic environment of the GC. These findings show an important role of lncRNAs in general and potentially also in the pathogenesis of HL.



***Exploring the potential of circulating free-DNA as biomarker and source for mutation detection***

Despite significant technological advances made in the past few years, many of them are still largely unexplored in primary HL samples, because of the scarcity of the malignant HRS cells in diagnostic biopsies. Some studies successfully enriched primary HRS cells using laser capture microdissection8,[52-55] or FACS[1]. Although this is doable, it is still challenging to obtain sufficient amounts of starting material for NGS approaches. Recently, a non-invasive prenatal test in a pregnant woman revealed genomic imbalances in the circulating cell-free DNA (cfDNA) fraction consistent with copy number aberrations observed in HL. Subsequently, she was diagnosed with early-stage nodular sclerosis HL during gestation.[56] Based on this study, genomic

imbalances present in HRS cells can be identified by massive parallel sequencing of cfDNA. Increased levels of cfDNA have been reported in HL patients.[57] Potentially we can also screen for presence of somatic mutations in the cfDNA fraction, which might be more straightforward as compared to purifying primary HRS cells. From the results presented in this thesis, we already have a list of genes with recurrent mutations and we could combine this with recurrent mutations found in other studies. It is interesting to check whether these mutations are detectable in cfDNA fractions of HL patients.

***Functional studies and downstream target genes***

Genomic data alone are normally insufficient to proof that a candidate gene is responsible for the observed phenotypes. To further explore the functionality of selected candidates (CD58, CSF2RB and MYB), knockdown and overexpression of wild type and mutant variants in HL cell lines should be done. In chapter 5 of this thesis, we used a shRNA approach to knockdown the MYB gene. In one of the cell lines, we saw no phenotype related to cell growth, while in the three other cell lines tested we did see a phenotype. Maybe the knockdown efficiency was not sufficient in the unresponsive cell line. We could change to a CRISPR/Cas9 based knock out approach. Recently, the bacterial clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated (Cas) system, which is known as the bacterial adaptive immune system that confers resistance against bacteriophages, was demonstrated as an efficient gene-targeting technology with the potential for multiplexed genome editing.[58-62] The ease and high efficiency of the CRISPR/Cas9 system makes it widely applicable.[63] For a more in depth understanding of the role of MYB, we should also perform the microarray analysis for the cell lines with truncated MYB protein to establish putative differences with the MYB wild type HL cell lines. The truncated MYB might have different target genes compared to wild type MYB.

Functional studies in HL cell lines may not represent a good model for the tumor cells *in vivo* since in the cell lines, the tumor cells have become independent of the reactive infiltrate. Unfortunately there is no good HL animal model that can mimic the situation of a minority of tumor cells surrounded by a predominant reactive infiltrate. In this thesis, we found that loss of CD58 expression in tumor cells is possibly a late event associated with relapse of disease. It would be interesting to establish an *in vivo* monitor of tumor relapse in HL. Maybe in the future, a PDX model of HL can be established, similar to the recently established PDX models of DLBCL.[64]

## REFERENCES


1. Reichel J, Chadburn A, Rubinstein PG, Giulino-Roth L, Tam W, Liu Y et al. Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. Blood 2015; 125(7): 1061-72.
2. Liu Y, Abdul Razak FR, Terpstra M, Chan FC, Saber A, Nijland M et al. The mutational landscape of Hodgkin lymphoma cell lines determined by whole-exome sequencing. Leukemia 2014; 28(11): 2248-51.
3. Nomoto J, Hiramoto N, Kato M, Sanada M, Maeshima AM, Taniguchi H et al. Deletion of the TNFAIP3/A20 gene detected by FICTION analysis in classical Hodgkin lymphoma. BMC Cancer 2012; 12: 457.
4. Pizzi M, Boi M, Bertoni F, Inghirami G. Emerging therapies provide new opportunities to reshape the multifaceted interactions between the immune system and lymphoma cells. Leukemia 2016.
5. Diepstra A, Niens M, Vellenga E, van Imhoff GW, Nolte IM, Schaapveld M et al. Association with HLA class I in Epstein-Barr-virus-positive and with HLA class III in Epstein-Barr-virus-negative Hodgkin's lymphoma. Lancet 2005; 365(9478): 2216-24.
6. Murray PG, Constandinou CM, Crocker J, Young LS, Ambinder RF. Analysis of major histocompatibility complex class I, TAP expression, and LMP2 epitope sequence in Epstein-Barr virus-positive Hodgkin's disease. Blood 1998; 92(7): 2477-83.
7. Oudejans JJ, Jiwa NM, Kummer JA, Horstman A, Vos W, Baak JP et al. Analysis of major histocompatibility complex class I expression on Reed-Sternberg cells in relation to the cytotoxic T-cell response in Epstein-Barr virus-positive and -negative Hodgkin's disease. Blood 1996; 87(9): 3844-51.
8. Steidl C, Shah SP, Woolcock BW, Rui L, Kawahara M, Farinha P et al. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. Nature 2011; 471(7338): 377-81.
9. Diepstra A, van Imhoff GW, Karim-Kos HE, van den Berg A, te Meerman GJ, Niens M et al. HLA class II expression by Hodgkin Reed-Sternberg cells is an independent prognostic factor in classical Hodgkin's lymphoma. J Clin Oncol 2007; 25(21): 3101-8.
10. Moreno CS, Beresford GW, Louis-Plence P, Morris AC, Boss JM. CREB regulates MHC class II expression in a CIITA-dependent manner. Immunity 1999; 10(2): 143-51.
11. Green MR, Kihira S, Liu CL, Nair RV, Salari R, Gentles AJ et al. Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. Proc Natl Acad Sci U S A 2015; 112(10): E1116-25.
12. Schneider M, Schneider S, Zuhlke-Jenisch R, Klapper W, Sundstrom C, Hartmann S et al. Alterations of the CD58 gene in classical Hodgkin lymphoma. Genes Chromosomes Cancer 2015; 54(10): 638-45.
13. Challa-Malladi M, Lieu YK, Califano O, Holmes AB, Bhagat G, Murty VV et al. Combined genetic inactivation of beta2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. Cancer Cell 2011; 20(6): 728-40.
14. Srivastava SK, Bhardwaj A, Singh S, Arora S, McClellan S, Grizzle WE et al. Myb overexpression overrides androgen depletion-induced cell cycle arrest and apoptosis in prostate cancer cells, and confers aggressive malignant traits: potential role in castration resistance. Carcinogenesis 2012; 33(6): 1149-57.
15. Thorner AR, Parker JS, Hoadley KA, Perou CM. Potential tumor suppressor role for the c-Myb oncogene in luminal breast cancer. PLoS One 2010; 5(10): e13073.

**16.** Biroccio A, Benassi B, D'Agnano I, D'Angelo C, Buglioni S, Mottolese M et al. c-Myb and Bcl-x overexpression predicts poor prognosis in colorectal cancer: clinical and experimental findings. Am J Pathol 2001; 158(4): 1289-99.

**17.** Uttarkar S, Dasse E, Coulibaly A, Steinmann S, Jakobs A, Schomburg C et al. Targeting acute myeloid leukemia with a small molecule inhibitor of the Myb/p300 interaction. Blood 2016; 127(9): 1173-82.

**18.** Barletta C, Lazzaro D, Prosperi Porta R, Testa U, Grignani F, Ragusa RM et al. C-MYB activation and the pathogenesis of ovarian cancer. Eur J Gynaecol Oncol 1992; 13(1): 53-9.

**19.** Shi J, Whyte WA, Zepeda-Mendoza CJ, Milazzo JP, Shen C, Roe JS et al. Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. Genes Dev 2013; 27(24): 2648-62.

**20.** Love C, Sun Z, Jima D, Li G, Zhang J, Miles R et al. The genetic landscape of mutations in Burkitt lymphoma. Nat Genet 2012; 44(12): 1321-5.

**21.** Wright JB, Brown SJ, Cole MD. Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells. Mol Cell Biol 2010; 30(6): 1411-20.

**22.** Persson M, Andren Y, Mark J, Horlings HM, Persson F, Stenman G. Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. Proc Natl Acad Sci U S A 2009; 106(44): 18740-4.

**23.** Zang ZJ, Cutcutache I, Poon SL, Zhang SL, McPherson JR, Tao J et al. Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes. Nat Genet 2012; 44(5): 570-4.

**24.** Qi C, Zhu YT, Hu L, Zhu YJ. Identification of Fat4 as a candidate tumor suppressor gene in breast cancers. International journal of cancer 2009; 124(4): 793-8.

**25.** Zhang J, Jima D, Moffitt AB, Liu Q, Czader M, Hsi ED et al. The genomic landscape of mantle cell lymphoma is related to the epigenetically determined chromatin state of normal B cells. Blood 2014; 123(19): 2988-96.

**26.** Parry M, Rose-Zerilli MJ, Gibson J, Ennis S, Walewska R, Forster J et al. Whole exome sequencing identifies novel recurrently mutated genes in patients with splenic marginal zone lymphoma. PLoS One 2013; 8(12): e83244.

**27.** Morin RD, Mungall K, Pleasance E, Mungall AJ, Goya R, Huff RD et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. Blood 2013; 122(7): 1256-65.

**28.** Fujimoto K, Shibasaki T, Yokoi N, Kashima Y, Matsumoto M, Sasaki T et al. Piccolo, a Ca2+ sensor in pancreatic beta-cells. Involvement of cAMP-GEFII.Rim2. Piccolo complex in cAMP-dependent exocytosis. J Biol Chem 2002; 277(52): 50497-502.

**29.** Lohr JG, Stojanov P, Lawrence MS, Auclair D, Chapuy B, Sougnez C et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc Natl Acad Sci U S A 2012; 109(10): 3879-84.

**30.** Fabbri G, Khiabanian H, Holmes AB, Wang J, Messina M, Mullighan CG et al. Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. J Exp Med 2013; 210(11): 2273-88.

**31.** Zhang ZC, Zheng LQ, Pan LJ, Guo JX, Yang GS. ZNF217 is overexpressed and enhances cell migration and invasion in colorectal carcinoma. Asian Pac J Cancer Prev 2015; 16(6): 2459-63.

**32.** Shida A, Fujioka S, Kurihara H, Ishibashi Y, Mitsumori N, Omura N et al. Prognostic significance of ZNF217 expression in gastric carcinoma. Anticancer Res 2014; 34(9): 4813-7.

**33.** ZNF217 upregulation in colorectal cancer and its oncogenicity. PLoS One 2015; 10(1): e0116170.

**34.** Westernberg L, Conche C, Huang YH, Rigaud S, Deng Y, Siegemund S et al. Non-canonical antagonism of PI3K by the kinase Itpkb delays thymocyte beta-selection and renders it Notch-dependent. Elife 2016; 5.

**35.** Dubois S, Viailly PJ, Mareschal S, Bohers E, Bertrand P, Ruminy P et al. Next-Generation Sequencing in Diffuse Large B-Cell Lymphoma Highlights Molecular Divergence and Therapeutic Opportunities: a LYSA Study. Clin Cancer Res 2016; 22(12): 2919-28.

**36.** Urayama KY, Jarrett RF, Hjalgrim H, Diepstra A, Kamatani Y, Chabrier A et al. Genome-wide association study of classical Hodgkin lymphoma and Epstein-Barr virus status-defined subgroups. J Natl Cancer Inst 2012; 104(3): 240-53.

**37.** Twa DD, Chan FC, Ben-Neriah S, Woolcock BW, Mottok A, Tan KL et al. Genomic rearrangements involving programmed death ligands are recurrent in primary mediastinal large B-cell lymphoma. Blood 2014; 123(13): 2062-5.

**38.** Van Roosbroeck K, Ferreiro JF, Tousseyn T, van der Krogt JA, Michaux L, Pienkowska-Grela B et al. Genomic alterations of the JAK2 and PDL loci occur in a broad spectrum of lymphoid malignancies. Genes Chromosomes Cancer 2016; 55(5): 428-41.

**39.** Plattel W KJ, Diepstra A, Visser L, van den Berg A. The role of microRNAs in Hodgkin's lymphoma. In: CH L (ed) Micro-RNAs in Medicine. John Wiley & Son, Inc, 2013, p pp 435e447.

**40.** Sandhu SK, Croce CM, Garzon R. Micro-RNA Expression and Function in Lymphomas. Adv Hematol 2011; 2011: 347137.

**41.** Van Vlierberghe P, De Weer A, Mestdagh P, Feys T, De Preter K, De Paepe P et al. Comparison of miRNA profiles of microdissected Hodgkin/Reed-Sternberg cells and Hodgkin cell lines versus CD77+ B-cells reveals a distinct subset of differentially expressed miRNAs. Br J Haematol 2009; 147(5): 686-90.

**42.** Gibcus JH, Tan LP, Harms G, Schakel RN, de Jong D, Blokzijl T et al. Hodgkin lymphoma cell lines are characterized by a specific miRNA expression profile. Neoplasia 2009; 11(2): 167-76.

**43.** Navarro A, Gaya A, Martinez A, Urbano-Ispizua A, Pons A, Balague O et al. MicroRNA expression profiling in classic Hodgkin lymphoma. Blood 2008; 111(5): 2825-32.

**44.** Navarro A, Diaz T, Cordeiro A, Beya MD, Ferrer G, Fuster D et al. Epigenetic regulation of microRNA expression in Hodgkin lymphoma. Leuk Lymphoma 2015; 56(9): 2683-9.

**45.** Jones K, Nourse JP, Keane C, Bhatnagar A, Gandhi MK. Plasma microRNA are disease response biomarkers in classical Hodgkin lymphoma. Clin Cancer Res 2014; 20(1): 253-64.

**46.** Spizzo R, Almeida MI, Colombatti A, Calin GA. Long non-coding RNAs and cancer: a new frontier of translational research? Oncogene 2012; 31(43): 4577-87.

**47.** Prensner JR, Chinnaiyan AM. The emergence of lncRNAs in cancer biology. Cancer Discov 2011; 1(5): 391-407.

**48.** Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y et al. The landscape of long noncoding RNAs in the human transcriptome. Nat Genet 2015; 47(3): 199-208.

**49.** White NM, Cabanski CR, Silva-Fisher JM, Dang HX, Govindan R, Maher CA. Transcriptome sequencing reveals altered long intergenic non-coding RNAs in lung cancer. Genome Biol 2014; 15(8): 429.

**50.** Verma A, Jiang Y, Du W, Fairchild L, Melnick A, Elemento O. Transcriptome sequencing reveals thousands of novel long non-coding RNAs in B cell lymphoma. Genome Med 2015; 7: 110.
**51.** Tayari M, Winkle M, Kortman G, Sietzema J, de Jong D, Terpstra M et al. Long Noncoding RNA Expression Profiling in Normal B-Cell Subsets and Hodgkin Lymphoma Reveals Hodgkin and Reed-Sternberg Cell-Specific Long Noncoding RNAs. Am J Pathol 2016.
**52.** Bargou RC, Leng C, Krappmann D, Emmerich F, Mapara MY, Bommert K et al. High-level nuclear NF-kappa B and Oct-2 is a common feature of cultured Hodgkin/Reed-Sternberg cells. Blood 1996; 87(10): 4340-7.
**53.** Lake A, Shield LA, Cordano P, Chui DT, Osborne J, Crae S et al. Mutations of NFKBIA, encoding IkappaB alpha, are a recurrent finding in classical Hodgkin lymphoma but are not a unifying feature of non-EBV-associated cases. Int J Cancer 2009; 125(6): 1334-42.
**54.** Joos S, Menz CK, Wrobel G, Siebert R, Gesk S, Ohl S et al. Classical Hodgkin lymphoma is characterized by recurrent copy number gains of the short arm of chromosome 2. Blood 2002; 99(4): 1381-7.
**55.** Hartmann S, Martin-Subero JI, Gesk S, Husken J, Giefing M, Nagel I et al. Detection of genomic imbalances in microdissected Hodgkin and Reed-Sternberg cells of classical Hodgkin's lymphoma by array-based comparative genomic hybridization. Haematologica 2008; 93(9): 1318-26.
**56.** Vandenberghe P, Wlodarska I, Tousseyn T, Dehaspe L, Dierickx D, Verheecke M et al. Non-invasive detection of genomic imbalances in Hodgkin/Reed-Sternberg cells in early and advanced stage Hodgkin's lymphoma by sequencing of circulating cell-free DNA: a technical proof-of-principle study. Lancet Haematol 2015; 2(2): e55-65.
**57.** Hohaus S, Giachelia M, Massini G, Mansueto G, Vannata B, Bozzoli V et al. Cell-free circulating DNA in Hodgkin's and non-Hodgkin's lymphomas. Ann Oncol 2009; 20(8): 1408-13.
**58.** Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N et al. Multiplex genome engineering using CRISPR/Cas systems. Science 2013; 339(6121): 819-23.
**59.** Malina A, Mills JR, Cencic R, Yan Y, Fraser J, Schippers LM et al. Repurposing CRISPR/Cas9 for in situ functional assays. Genes Dev 2013; 27(23): 2602-14.
**60.** Wang Y, Li Z, Xu J, Zeng B, Ling L, You L et al. The CRISPR/Cas system mediates efficient genome engineering in Bombyx mori. Cell Res 2013; 23(12): 1414-6.
**61.** Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science 2014; 343(6166): 80-4.
**62.** Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 2013; 153(4): 910-8.
**63.** Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 2014; 32(4): 347-55.
**64.** Chapuy B, Cheng H, Watahiki A, Ducar MD, Tan Y, Chen L et al. Diffuse large B-cell lymphoma patient-derived xenograft models capture the molecular and biological heterogeneity of the disease. Blood 2016; 127(18): 2203-13.

# Appendices

## NEDERLANDSE SAMENVATTING

Het Hodgkin lymfoom (HL) heeft een karakteristieke histologie met een minderheid aan tumor cellen, de zogenaamde Hodgkin Reed-Sternberg (HRS) cellen, en een grote overmaat aan immuun cellen. We onderscheiden twee varianten: het klassieke (c)HL en nodulair lymfocyt predominant NLPHL. De interactie tussen de tumorcellen en de cellen in de omgeving is belangrijk voor de overleving van de tumor cellen. De Hodgkin lymfoom tumor cellen worden gekenmerkt door constitutieve activatie van de NF-KB en JAK/STAT signaleringsroutes. Dit karakteristieke fenotype wordt o.a. veroorzaakt door mutaties in genen die activatie van deze signaleringsroutes reguleren. Deze gen mutaties zijn in eerste instantie aangetoond in HL cellijnen en later bevestigd in tumor cellen geïsoleerd uit primaire Hodgkin lymfoom weefsels met behulp van laser microdissectie technieken. In deze initiële studies werd er steeds een a priori selectie gemaakt van kandidaat genen gerelateerd aan de reeds bekende signaleringsroutes.

In dit promotie onderzoek, hebben we de eerste genoom brede mutatie analyse studie gedaan in HL cellijnen en vervolgens hebben we de functionele consequenties van een aantal van de gevonden gen mutaties onderzocht in HL.

In **hoofdstuk 2** hebben we het volledige coderende deel van het humane genoom onderzocht op gen mutaties in 7 HL cellijnen. We vonden in totaal 463 genen die frequent gemuteerd waren in de HL cellijnen ongeacht het HL subtype en 373 genen die frequent gemuteerd waren specifiek in de 6 cHL cellijnen. Gen ontologie analyse van deze genen sets toonde aan dat cel adhesie en cel ontwikkeling en differentiatie de meest voorkomende gen functie categorieën waren. 44 van de 373 genen met mutaties in cHL hadden ook een significant verschil in expressie tussen de tumorcellen en het celtype waar cHL uit ontstaat, namelijk normale kiemcentrum B cellen. Chromosomale kopie veranderingen werden gevonden voor een klein deel van de gemuteerde genen.

De genen waarvan reeds bekend was dat ze frequent gemuteerd zijn in HL, werden ook gevonden in onze studie. Daarnaast vonden we ook een goede overlap met sequentie analyse data van ruim 1600 genen beschikbaar op de COSMIC website voor deze cellijnen. Voor een deel van de mutaties hebben we validatie experimenten gedaan aan de hand van RNA-seq data en aan de hand van Sanger-sequentie analyse van geselecteerde mutaties. Dit geeft aan dat onze analyse betrouwbare data heeft opgeleverd.

Naast de reeds bekende mutaties vonden we ook mutaties in een aantal immuunrespons geassocieerde genen, o.a. *B2M, CIITA, CD58, en CSF2RB*. Verder vonden we ook mutaties in de transcriptie factor MYB, een eiwit dat ook oncogene eigenschappen heeft in meerdere soorten tumoren. Deze studie heeft aldus een aantal nieuwe genen gevonden die vaak gemuteerd zijn in HL en die mogelijk bijdragen aan de pathogenese van HL.

In **hoofdstuk 3**, hebben we de CD58 mutaties en expressie verder onderzocht in HL. CD58 (LFA-3) is een adhesie molecuul betrokken bij immuun herkenning van tumor cellen door te binden aan de CD2 receptor op cytotoxische T cellen. In diffuus groot cellige B cel lymfomen werden ook mutaties in het CD58 gen gevonden in eerdere studies en werd aangetoond dat dit gen betrokken is bij de effectiviteit van de anti-tumor immuun respons. We konden de aanwezigheid van mutaties in alle drie HL cellijnen bevestigen op DNA en RNA niveau. Expressie analyse van mRNA en eiwit gaf aan dat de expressie in de cellijnen met een mutatie erg laag of zelfs afwezig was, terwijl in cellijnen met wild type sequentie er wel een normale expressie werd gevonden. In primair weefsel zagen we geen verlies van CD58 expressie in patiënten die goed respondeerden op de behandeling, terwijl we in 11% van de patiënten die niet goed reageerden op de therapie verlies zagen van CD58 expressie. Dit verlies is mogelijk een late gebeurtenis aangezien we dit alleen vonden in weefsel van patiënten waarbij de tumor na afloop van de behandeling weer terug kwam. Verder zijn ook de cellijnen afkomstig van patiënten die slecht hebben gereageerd op de therapie, wat consistent is met onze bevindingen op patiënten weefsel. Deze resultaten geven aan dat verlies van CD58 mogelijk bijdraagt aan de ontsnapping van de HRS cellen aan immuun responsen in een laat stadium van de ziekte.

In **hoofdstuk 4**, hebben we de relevantie van mutaties in het *CSF2RB* gen verder onderzocht. *CSF2RB* codeert de β keten van de receptor van interleukin-3 (IL-3), IL-5 en granulocytic macrophage colony-stimulating factor (GM-CSF). In tegenstelling tot de β keten zijn de α ketens voor deze drie cytokines specifiek. Binding van deze cytokines aan de receptor resulteert in signalering van JAK2/STAT5 en dit geeft overleving, cel groei en cel differentiatie signalen. We konden de zes mutaties die we hadden aangetoond in vier HL cellijnen allemaal bevestigen op DNA en RNA niveau. Door de relatief grote afstand tussen de gemuteerde posities in twee van de vier cellijnen konden we niet aantonen of deze op hetzelfde of verschillende gen kopieën zaten. In een recent gepubliceerde studie werden mutaties in CSF2RB ook aangetoond in HRS cellen van primaire weefsel biopten. We zagen geen duidelijk verschillen tussen eiwit expressie niveau in gemuteerde en wild type

cellijnen. Behandeling van de cellijnen met de drie cytokines die kunnen binden aan de receptor lieten geen duidelijke effecten zien op groei van de HL cellijnen, ongeacht de aanwezigheid van mutaties. Dit zou mogelijk verklaart kunnen worden door activerende mutaties in de JAK/STAT signaleringsroutes. Verder onderzoek naar de consequentie van deze mutaties in de pathogenese van HL is nog nodig.

In **hoofdstuk 5** hebben we de rol van mutaties in het *MYB* gen verder onderzocht. MYB is een transcriptie activator die betrokken is bij celdeling en differentiatie van hematopoëtische stamcellen. We konden de twee frameshift mutaties bevestigen op DNA en RNA niveau. Op eiwit niveau zagen we zoals verwacht geen aankleuring in de twee gemuteerde HL cellijnen met een antilichaam tegen het uiteinde van het eiwit. Met antilichamen gericht tegen het begin van het eiwit zagen we een kleiner eiwit product, overeenkomstig met de gevonden mutaties. In de cellijnen met wild type MYB zagen we het eiwit met normale lengte. In tegenstelling tot de HL cellijnen, zagen we in HRS cellen van primaire weefselbiopten geen aankleuring.

Inhibitie van MYB in HL cellijnen had een sterk negatief effect op de groei, ongeacht aanwezigheid van een mutatie. Dit negatieve effect werd veroorzaakt door een toename van het aantal apoptotische cellen. Inhibitie van MYB had geen effect op de expressie van de oncogene transcriptie factor MYC. Vice versa had inhibitie van MYC wel een effect op de eiwit levels van MYB. Genexpressie analyses van HL cellijnen met controle en MYB eiwit verlagende shRNA constructen toonde dat 60 genen een significant veranderde expressie hadden. Een van de genen die was veranderd in beide geteste cellijnen was het anti-apoptotische BCL2 gen. Gen ontologie analyses van de 60 genen toonde aan dat MYB gereguleerde genen functies hebben gerelateerd aan proliferatie, apoptose, cel adhesie en immuunrespons. Deze processen zijn allemaal relevant voor de pathogenese van HL. Verder onderzoek moet aantonen welk van deze genen bijdragen aan overleving van HRS cellen en of aanwezigheid van mutaties in het MYB gen het oncogene effect nog verder kan versterken.

Samenvattend hebben we in dit promotie onderzoek een genoom brede mutatie analyse gedaan van alle eiwit coderende genen in HL. We konden de reeds bekende gen mutaties valideren en toonden tevens aan dat mutaties in immuun gerelateerde genen ook frequent voorkomen in HL. Deze mutaties dragen mogelijk bij aan de overleving van de tumor cellen en aan de ontsnapping aan effectieve anti-tumor immuunresponsen.

## AUTHORS' AFFILIATIONS

Fazlyn Reeny Abdul Razak[1]
Anke van den Berg[1]
Arjan Diepstra[1]
Lydia Visser[1]
Joost Kluiver[1]
Klass Kok[2]
Martijn Terpstra[2]
Yuxuan Liu[1]
Ali Saber[1]
Marcel Nijland[3]
Gustaaf van Imhoff[3]
Christian Steidl[4]
Fong Chun Chan[4]
Randy Gascoyne[4]
Debora de Jong[1]
Jasper Koerts[1]
Bea Rutgers[1]
Emmie de Vries[1]
Michelle Oomkens[1]

[1]Department of Pathology & Medical Biology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.

[2]Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.

[3]Department of Hematology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands.

[4]Centre for Lymphoid Cancer, Department of Experimental Therapeutics, Vancouver, Canada.

## AUTHOR BIOGRAPHY

Fazlyn Reeny Abdul Razak was born on 22nd May 1985 in Terengganu, Malaysia. In August 2007, she obtained her Bachelor of Biomedical Science degree from The National University of Malaysia. After receiving her bachelor's degree, she worked at the Department of Biochemistry, The National University of Malaysia as a research assistant (under the supervision of Prof. Wan Zurinah Wan Ngah) on alternative medicine oncology. The topic raised her interest for research and made her decide to proceed her master degree. She got her Master of Medical Science in Biochemistry from The National University of Malaysia in 2013. During her master studies, she also working as a researcher in the Molecular Pathology Unit, Cancer Research Center, Institute for Medical Research (IMR) in Kuala Lumpur. In December 2013, she started her PhD project under supervision of Prof. Anke van den Berg and Dr. Arjan Diepstra at University of Groningen, University Medical Center Groningen, The Netherlands. She is currently working as a researcher in the Molecular Pathology Unit, Cancer Research Center, Institute for Medical Research (IMR) in Kuala Lumpur.

**ACKNOWLEDGEMENTS**

I am very thankful to Allah S.W.T for His blessings that enables me to complete my PhD journey. Finally I finished my PhD thesis, and I would like to thank numerous people for their help and support during my PhD studies. Many people have contributed to the research in their own particular way. It is not easy to summarize my all acknowledgments in a few words, but just I can say that I am really grateful to all people I met during my PhD.

First and foremost, I am extremely grateful to my promoter Prof. Anke van den Berg. Dear Anke, thank you for accepting me as a PhD student and to be part of this group. Thank you for your enthusiasm, guidance, and unrelenting support throughout my PhD. I admire how you can do so much so quickly and always think logical. Your mix of straightforward criticism combined with heart-warming support have given me great confidence as a researcher. I would like to thank Dr. Arjan Diepstra for being my co-promoter. Dear Arjan, thank you for all the fruitful discussions in general and on the methodological approach especially related to the histopathology. Thank you also for your easy-going attitude and always tried to encourage me and keep me motivated about what I was doing. Anke and Arjan, I have been extremely lucky to have both of you who cared so much about my work, who responded to my questions and queries so promptly and making sure the research continued to progress towards a finished thesis. Your frequent insights and patience with me are always appreciated. I owe a great debt of gratitude to both of you.

I also want to take a moment to thank the committee members, Prof. Gerwin Huls, Prof. Steven Jong and Prof. Paul Murray. Thank you for investing time and I feel proud and honored that you have accepted to be on my committee.

A very special thanks to Dr. Lydia Visser and Dr. Joost Kluiver. Dear Lydia, thank you for always be there for me during my stay in Groningen and for cheering me up when I needed it. For sure I will miss your delicious cakes. Joost, thank you very much for helping me during these years and thank you for always asking me, “How are you doing? Everything OK?” I also like to thanks Dr. Tan Lu ping for being a positive influence, for giving me the confidence to pursue my PhD abroad and for recommending me to apply for the PhD study. Prof. Philin Kluin, I appreciate your stimulating questions during my presentations.

I would like to thank you people from Department of Genetics, Dr. Klaas Kok and Martijn Terpstra. Dear Klaas, your knowledge in next generation

sequencing techniques is highly appreciated. Dear Martijn, thank you for helping me with the bioinformatic analysis for my projects. I would like to thank all the people in the genetics department sequencing facility.

I would also like to thank you people from Department of Hematology, Dr. Gustaaf van Imhoff, Prof. Hanneke Kluin-Nelemans, Dr. Tom van Meerten and Dr. Marcel Nijland. Dear Marcel, you gave me some good pieces of advice related to my studies and life.

I am also thankful to the co-authors and collaborators for contributing their work and time. People from Department of Experimental Therapeutics, Vancouver: Christian Steidl, Fong Chun Chan and Randy Gascoyne.

I would like to thank my fellow friends for providing a sense of community and camaraderie. Thank you for accompanying me on this adventure and providing many laughs, exciting discussions and a fun during my PhD studies. So a big thank you to all of you; Ali, Rianne, Emine, Yuan, Melanie, Mina, Ahmad, Mathilde, Agnieszka, Fubiao, Jennie, Hataitip, Karolin, Jan, Wouter, Lotte, Geok Wee, Ferronika, Yuxuan, Kushi, Nato, Rae, Juan, Ben, Cheng, Siobhan, Rik and Joost. I'm so blessed to have some really memorable friends. An extra thank you to Nato, Rae, Mina and Ali, the ones I tend to turn to first when I have something to discuss, both in times of problems and of victories. Or just need a break.

Big thanks to my two Dutch students, Emmie and Michelle. It was a great chance to have both of you beside me during my research. Being able to discuss my work with you has been invaluable for developing my ideas.

I would like to extend great thanks to all the people and technicians in the molecular diagnostic, DNA and O&O lab for letting me have this unforgettable experience working in the lab, especially: Debora, Jasper, Mirjam, Lorian, Sippie, Klaas, Geert, Roland, Jesper, Annika, Bart, Bea, Tineke, Weird, Marian, Monique, Hans, Jantine, Ingrid, Martin, Inge, Erik, Rianne, Lisette, Paskal, Annelies and Marco. Dear Debora and Jasper, both of you were a huge help in the lab. Thank you for patiently answered all my beginners' questions and were so kind to me. Dear Bea, thank you for helping me whenever I had doubts. Tineke, thank you for helping me with the sectioning of tissue blocks. Weird, I really appreciated your help with the LDM. Dear Annika, Jantine and Ingrid, thank you for all the nice chats in the labs.

I am grateful to all the administrative staff of the department of Pathology for practical support throughout my PhD. I also would like to thank our histology lab and archive for their magnificent cooperation and support. Thank you Geert, Henk and Theo for your patience and help with the FACS. I would like to thank the GSMS and University of Groningen not only for providing the funding which allowed me to undertake this research, but also for giving me the opportunity to attend conferences and meet so many interesting people.

To my fellow friends I've met in Groningen (Nadiah, Zetty & Abang Amir, Iqa & Syafiq, Azie, Shak, Julie, Ziad, Khalisa & Alex, Najib & Kak Anis, Zul, EeSoo, Kak Ili & Yusof, Mai, Ina, Eina and Gulbahar's family). You were like a family. Thank you for always being helpful and offered me food, advices, laughter and more. I really appreciate the time we spent together for serious discussions, advices, jokes and nonsense talking. I have enjoyed the friendship and support. Special thanks to Nadiah, Zetty, Iqa, Azie and Shakila. Girls, thank you for showing me that there is always someone out there who cares and is willing to listen. Without your constant support, love, friendship and ability to make me laugh I couldn't finish it.

I am indebted to all my friends and colleagues who have supported me over the last few years: friends from SRKK, SESMA, UKM and IMR. Especially my dear Ctee, Shimah, Aimie, Pijie and Ajim, thank you for always being there for me.

Rianne and Nadiah, it is great that you are willing to be my paranymphs.

I would like to thank my family especially my parents (Abdul Razak & Wan Fuziah) whose love, motivations and prayers have continued to guide me. I wish to also thank my siblings (Rozliyana Fazliyati, Rozailan Yusainy, Farah Reeza and Muhammad Firdaoz), brother-in-law (Burhan), sister-in-law (Noriyahtuakma) and my nieces & nephews (Pesal, Eena, Apiq, Rehan, Nia, Eera, Zidan, Emir, Zariff, Riaz, Aqil, Esya & Qays) for being such a wonderful family. My family have been so supportive and have given me unconditional encouragement throughout; I would not have been able to complete a PhD without them. I hope I have made them proud.

Finally, thank you so much to everyone who has helped me along my student journey. So many people have been so kind. My memories in Groningen will always be in my mind and heart. Thank you all!